

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2004, 00:01:12 ; Search time 819 Seconds

(without alignments)
10446.729 Million cell updates/sec

Title: US-09-980-364-1

Perfect score: 2014

Sequence: 1 gttcatctctcttcttaag.....gttgataatttcagact 2014

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2014	100.0	2014	4 AAC91398	AAC91398 Brassica
2	1867	92.7	2011	4 AAC91399	AAC91399 Brassica
3	726	36.0	4873	4 AAC91400	AAC91400 Brassica
4	420.6	20.9	2463	7 ACC00803	ACC00803 Glycine m
5	413.6	20.5	1824	7 ACC00802	ACC00802 Glycine m
6	413.6	20.5	2598	7 ACC00801	ACC00801 Glycine m
7	399.2	19.8	1749	7 ACC93248	ACC93248 Cellulose
8	385.8	19.2	1990	7 ACC00806	ACC00806 Glycine m
9	377.4	18.7	1818	9 ADD30878	ADD30878 Plant yie
10	377.4	18.7	1818	9 ADD30822	ADD30822 Plant yie
11	377.4	18.7	1818	9 ADD30333	ADD30333 Plant yie
12	357.4	17.7	1935	7 ACC00848	ACC00848 Gossypium
13	352	17.5	2168	7 ACC00804	ACC00804 Glycine m
14	347.8	17.3	1290	7 ACC00805	ACC00805 Glycine m
15	347.8	17.3	1954	7 ACC00850	ACC00850 Glycine m
16	335.2	16.6	1878	7 ACC00846	ACC00846 Glycine m
17	334	16.6	1874	7 ACC00799	ACC00799 Glycine m
18	332.2	16.5	2052	9 ADD30980	ADD30980 Plant yie
19	330.6	16.4	1668	3 AAC43005	AAC43005 Arabidops
20	330.6	16.4	2148	3 AAC59220	AAC59220 cDNA enco
21	330.6	16.4	2148	7 ABX13437	ABX13437 A. thalia
22	330.2	16.4	1959	7 ACC00847	ACC00847 Catalpa s
23	327.6	16.3	2374	7 ACC00849	ACC00849 Zea mays

ALIGNMENTS

RESULT 1
AAC91398
ID AAC91398 standard; cDNA; 2014 BP.
XX
AC AAC91398;
XX
DT 19-MAR-2001 (first entry)
XX
DE Brassica napus BNM3A cDNA.
XX
KW Brassica napus microspore embryo; BNM; microspore embryogenesis; BNM3A;
XX asexually derived embryo production; increased regenerative capacity; ss.
XX
OS Brassica napus.
XX
PN EP1057891-A1.
XX
PD 06-DEC-2000.
XX
PF 02-JUN-1999; 99EP-00201745.
XX
PR 02-JUN-1999; 99EP-00201745.
XX
PA (CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;
PI Boutillier K;
XX
DR WPI: 2001-026720/04.
DR P-PSDB; AAB50872.
XX
PT Brassica napus BNM3 DNA sequences, useful for transforming host cells to
PT produce asexually derived embryos.
XX
PS Claim 4; Fig 2; 51pp; English.
XX
CC The present sequence is given in a specification relating to a Brassica
CC napus microspore embryo (BNM) gene obtained during the induction of
CC microspore embryogenesis. The isolated DNA molecule can be used in a
CC vector that is used to transform plant cells. The vector can be used in
CC methods to produce asexually derived embryos, for modifying the
CC regenerative capacity of a plant, and for producing an apomictic plant
XX
SQ Sequence 2014 BP; 596 A; 411 C; 462 G; 545 T; 0 U; 0 Other;

Abn84481 Soybean A
Add30559 Plant yie
Aaa59222 cDNA enco
Add30349 Plant yie
Acc00794 Oryza sat
Abn84482 Rice AINT
Acc00791 Zea mays
Abn84480 Soybean A
Acc00792 Zea mays
Abn84484 Rice AINT
Acc00796 Oryza sat
Acc00797 Glycine m
Acc00844 Ricinus c
Add31036 Plant yie
Ade31572 Plant yie
Acc00789 Zea mays
Acc00845 Glycine m
Aac5369 Arabidops
Abq81396 Arabidops
Abk65225 Arabidops
Adb31800 DNA enco
Adc46704 Thalecres

XX	19-MAR-2001 (first entry)	QY	GAGGAGGACGGTGTGGTGGCTCACTAGGACTTTTCGATGATAAAGACATGCTGAGAA	549
DT		Db		
XX	Brassica napus BN3B CDNA.	QY	GAGGAGGACGGTGTGGTGGCTCACTAGGACTTTTCGATGATAAAGACATGCTGAGAA	480
DE		Db		
XX	Brassica napus microspore embryo; BNM; microspore embryogenesis; BNM3B;	QY	ATCAACCCGTGGATAATGTTGATAATCAAGAAATGCAATGCTGCAAAAGGCTGTCCC	609
KW	asexually derived embryo production; increased regenerative capacity; ss.	Db		
XX	Brassica napus.	QY	ATCAACCCGTGGATAATGTTGATAATCAAGAAATGCAATGCTGCAAAAGGCTGTCCC	540
OS		Db		
XX	Brassica napus.	QY	TCTCAATGAACCTCATCTACTTCTTGTGATAACCAACTACAGCAGTAACAACCTTGTG	669
DN	EP1057891-A1.	Db		
XX	06-DEC-2000.	QY	TCTCAATGAACCTCATCTACTTCTTGTGATAACCAACTACAGCAGTAACAACCTTGTG	600
PD		Db		
XX	02-JUN-1999; 99BP-00201745.	QY	CCCAAGGGAGACTATTGATAGCTTGAAGCTACACCGAAGAACTATTGAGAGTT	729
XX	02-JUN-1999; 99BP-00201745.	Db		
XX	(CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.	QY	CCCAAGGGAGACTATTGATGATAGCTTGAAGCTACACCGAAGAACTATTGAGAGTT	660
PA	(MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.	Db		
XX	Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;	QY	TTGGACAGAGCAGCTCTATATACCCCGGTGTTACAAGGCATCGGTGACAGGAAGATG	789
PI	Boutillier K;	Db		
XX	WPI; 2001-026720/04.	QY	TTGGACAGAGCAGCTCTATATACCCCGGTGTTACAAGGCATCGGTGACAGGAAGATG	720
DR	P-PSDB; AAB50873.	Db		
XX	Brassica napus BN3 DNA sequences, useful for transforming host cells to	QY	AGGCACATTTATGGGATAATAGTTGTAAGAGAGGCGCAACCGCAAGAGAGACAAG	849
PT	produce asexually derived embryos.	Db		
XX	Claim 4; Fig 2; 51pp; English.	QY	AGGCACATTTATGGGATAATAGTTGTAAGAGAGGCGCAACCGCAAGAGAGACAAG	780
XX	The present sequence is given in a specification relating to a Brassica	QY	TTTATTTGGGAGTTATGACAAAGAAAGAGAGAGGCTAGGGCTTATGATTTAGCCGAC	909
CC	napus microspore embryo (BNM) gene obtained during the induction of	Db		
CC	microspore embryogenesis. The isolated DNA molecule can be used in a	QY	TTTATTTGGGAGTTATGACAAAGAAAGAGAGGCTAGGGCTTATGATTTAGCCGAC	840
CC	vector that is used to transform plant cells. The vector can be used in	Db		
CC	methods to produce asexually derived embryos, for modifying the	QY	TCAAGTATTTGGGAAACCACTACTACTTATCTCCCATGAGGGAATATGAGAAAGAGA	900
CC	regenerative capacity of a plant, and for producing an apomictic plant	Db		
XX	Sequence 2011 BP; 610 A; 395 C; 462 G; 544 T; 0 U; 0 Other;	QY	TCAAGTATTTGGGAAACCACTACTACTTATCTCCCATGAGGGAATATGAGAAAGAGA	900
SQ		Db		
QY	Query Match 92.7%; Score 1867; DB 4; Length 2011;	QY	TAGAGAGATCAAGCACATGACAGGCAAGAGTATGTTGCTCTACTGCGCAGGAAAGTA	1029
Db	Best Local Similarity 97.8%; Pred. No. 0;	Db		
QY	Matches 1903; Conservative 0; Mismatches 40; Indels 2; Gaps 1;	QY	TAGAGAGATCAAGCACATGACAGGCAAGAGTATGTTGCTCTACTGCGCAGGAAAGTA	960
Db		Db		
QY	70 TTCTTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTCTCTGAAGTCAATCAATAA	QY	GCACAACAAGAAAGCTCGAGAGGATACGACATTCGCGGCATCAAAATTCAGAGATTAA	1209
Db	1 TTCTTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTCTCTGAAGTCAATCAATAA	Db		
QY	130 ACTGGTTAGGCTTTTCTCTCTCTCTTATGAACAAATCAACATCGTAAGGAGCTTACT	QY	CGGAGTCACTAACTTCGACATGACAGATCAACGTTTAAAGCAATCTCGAAAGCCCTA	1269
Db	61 ACTGGTTAGGCTTTTCTCTCTCTCTTATGAACAAATCAACATCGTAAGGAGCTTCT	Db		
QY	190 CTTCCACACACAAACCGTGTAGATGTCGCGGAGAGTACTGTACGATCCGACCGCTG	QY	CGGAGTCACTAACTTCGACATGACAGATCAACGTTTAAAGCAATCTCGAAAGCCCTA	1200
Db	121 CTTCCACACACAAACCGTGTAGATGTCGCGGAGAGTACTGTACGATCCGACCGCTG	QY	GTCTTCTCTATTGTTAGCGCGCAAAACGCTCTCAAGGAGGCTAAACCGTCCGTTCCAA	1329
QY	250 CTTCCAGTGTCTTACGACCAATCAAAACATCGTTTCTCTCTCTCTCTCTCTCTCT	Db		
Db	181 CTTCCAGTGTCTTACGACCAATCAAAACATCGTTTCTCTCTCTCTCTCTCTCTCT	QY	GTCTTCTCTATTGTTAGCGCGCAAAACGCTCTCAAGGAGGCTAAACCGTCCGTTCCAA	1260
QY	310 ATGCTTTTACGAGACAAACATAGTCACTCCCGAGATGGGACATCAATGGTTGTCAT	QY	GTCTTCTCTATTGTTAGCGCGCAAAACGCTCTCAAGGAGGCTAAACCGTCCGTTCCAA	1389
Db	241 ATGCTTTTACGAGACAAACATAGTCACTCCCGAGATGGGACATCAATGGTTGTCAT	Db		
QY	370 GCAATACATCCAAACGATGAGCAAGATGACCAAGCTTGAGAAATTCCTTGGCGGCA	QY	GTCTTCTCTATTGTTAGCGCGCAAAACGCTCTCAAGGAGGCTAAACCGTCCGTTCCAA	1449
Db	301 GTAATACATCCAAACGATGAGCAAGATGACCAAGCTTGAGAAATTCCTTGGCGGCA	Db		
QY	430 CCACCAAGATTTACAAACCAACGATGAGCAAGATGACCAAGCTTGAGAAATTCCTTGG	QY	GTCTTCTCTATTGTTAGCGCGCAAAACGCTCTCAAGGAGGCTAAACCGTCCGTTCCAA	1509
Db	361 CCACCAAGATTTACAAACCAACGATGAGCAAGATGACCAAGCTTGAGAAATTCCTTGG	Db		
		QY	GTCTTCTCTATTGTTAGCGCGCAAAACGCTCTCAAGGAGGCTAAACCGTCCGTTCCAA	1629

QY 875 AGAARAGAGCTAGGCTTATGATTAGCGCAGCTCAAGTATGCGGAAACACCACTAC 934
DB 953 AGAARAGGAGCTAGAGCTTACGATTTGCGAGCACTAAATGCTGGGAAACACCAAC 1012
QY 935 TACTAATCTCCCATGAGCGAATATGAAAGAGGTGAGAGATGAAGCAATGACAAAG 994
DB 1013 AACAAATTTTCCAAATGAGCCACTATGAGAAAGAGTTGGAAGAAATGAAGCACTAG 1072
QY 995 GCGAGATATGTCCTCTAGCGGAGGAAAGTGTCTCTCGGTGCTCGATCGAT 1054
DB 1073 GCAAGAGTACGTGCGTCATTGAGAAGGAGTAGTGGGTTTTCGCGGTGCTATCAT 1132
QY 1055 TTATCGTGGAGTAACAGACATCACCAACATGGAAGATGCGAAGCTAGATAGGAGAGT 1114
DB 1133 TTATCGAGAGTACGAGACACCAACCAACATGGAAGGTGCGAAGGATTTGGAAGAT 1192
QY 1115 CGCGGTAAACAAGACTCTACTTGGAACTTTTGGACACACAGAGAGCTGCGAGGC 1174
DB 1193 TGTGCGCAACAGGATCTTTACTTGGAACTTTTACGCCCAAGAGAGCGAGCGAAGC 1252
QY 1175 ATACGACATTCGCGCCATCAAAATTCAGAGGATTAACCGGAGTCACTAACTTCGACATGAA 1234
DB 1253 ATATGATGAGCAGCAATCAAAATTCGAGGACTAAGTGTGTTTACAACTTTGACATGAG 1312
QY 1235 CAGATCAAGTTAAAGCAATCTCGAAAGCCCTAGTCTTCTCTATTTGGTAGCGCCGCAAA 1294
DB 1313 CAGATATGAGTGAAAGCATACTTGGAGACCACTTTGCCAATAGGTGGTGTGCAAA 1372
QY 1295 ACCTCTCAAGGA 1306
DB 1373 GCGTTTGAAGGA 1384

RESULT 7

AAC83248
ID AAC83248 standard; DNA; 1749 BP.
XX AAC83248;
AC AAC83248;
DT 14-MAR-2001 (first entry)
XX Cellulose synthase promoter sequence SEQ ID 3.
DE Cellulose synthase; lignin; secondary cell wall construction; wood pulp;
KW transgenic plant; paper manufacture; promoter; ds.
XX Unidentified.
OS WO200070058-A2.
XX 23-NOV-2000.
XX 17-MAY-2000; 2000WO-GB001890.
XX 18-MAY-1999; 99GB-00011379.
XX (UYMA-) UNIV VICTORIA MANCHESTER.
XX Turner S, Taylor N;
XX WPI; 2001-041015/05.
XX Cellulose synthase gene expressed during deposition of secondary cell
XX walls in lignin-containing cells; useful for modulating expression of
XX enzymes involved in synthesis of plant cell walls and to produce
XX transgenic plants.
XX Claim 10; Page 30-31; 49pp; English.

XX This invention relates to a cellulose synthase gene expressed during the
XX deposition of secondary cell walls in cells containing lignin. The
XX cellulose synthase gene is useful for regulating the expression of genes

CC specifically during secondary cell deposition in lignin containing cells.
CC it can be used to modify the structure and cellulose content of plant
CC secondary cell walls and to produce altered plant phenotypes specific to
CC the needs of a particular industry such as in reducing the lignin of wood
CC pulp for paper manufacturing. A construct containing a cellulose synthase
CC promoter sequence and a gene of interest may be used in a method for the
CC production of the product of the gene of interest in a host cell that
CC produces lignin, where the product is produced only during secondary cell
CC wall synthesis. The present sequence represents a cellulose synthase
CC promoter which can be used in the invention for the production of
CC transgenic plants expressing an exogenous gene during secondary cell wall
CC deposition in cells containing lignin
XX

XX Sequence 1749 BP; 551 A; 290 C; 354 G; 554 T; 0 U; 0 Other;

Query Match 19.8%; Score 399.2; DB 4; Length 1749;
Best Local Similarity 77.8%; Pred. No. 7.1e-92;
Matches 502; Conservative 0; Mismatches 123; Indels 49; Gaps 8;

QY 1257 CTCGAAGCCCTAGTCTTCTTATTTGGTAGCGCCCAAAAGCTCTCAAGGAGGCTAACCGT 1316
DB 1 CTCGAGAGCCCGAGTCTACCTATTGGTAGTTCTCGAAACGCTCTCAAGGAGCTTAAACAAT 60
QY 1317 CCGGTTCCAACTATGATGATCAGTAATAACGTTTCAGAGAGTGAGAAATAGTGCTAGC 1376
DB 61 CCGGTTCCAGCT---ATGATGATTTAGTAATAACGTTTCAGAGAGTGCAATATGTTAGC 117
QY 1377 GGTGGCAAAAACGCTGCGGTTTCAGATCATCAGGAGTAGATTTGAGCTTATTCGACCAA 1436
DB 118 GGTGGCAAAAACACTGCGGTTTCAG---CATCAGGGAATGGATTTGAGCTTATTCGACCAA 174
QY 1437 CATCAAGAGAGGTACAATGGTTATTATACAAATGAGGAAACTTGTCTCGGAGAGTCT 1496
DB 175 CAGCAGGAGAGGTAGTTGG---TTATTAACAATGAGGAAACTTGTCTACCGAGAGTACT 231
QY 1497 AGGGCTTGTTCACAAAGAGAGATGATCAACACCATTTCTTTGAGCAACACCGAGGCTC 1556
DB 232 AGGGTTTGTTCACAAAGAGAGAGGAAACAACACACTTCTTTGAGAAACTCGCCGAGTCA 291
QY 1557 ATGACTAATATCGATCATCAAGTCTGTTTCGGATGATCGGTACTGTTGTTGGAAT 1616
DB 292 ATGACTAATGTTGATCATCATGCTCTACCTCTGATGATTTCTGTTACCGGTTTGGAAAT 351
QY 1617 GTTGTGCTTATGTTGGTTTATCAAGGATTTGAGCC---CGGTTAACTGC 1664
DB 352 GTTGTGCTTATGTTGGTTTATCAAGGATTTGCAATCCCTGTTGGAACATCGTTAATTAC 411
QY 1665 GATGCTAGCGCTAGTGTAGTTGATTAACCGAAGAAACCATTTACTTTGCTCAG 1724
DB 412 GATCCCTTTACTGCTGTGAGATTGCTTCAACGCAAGAAATCATTTACTATGCTCAG 471
QY 1725 CAG-----CAGCAGACCCAGCGAGTGGAGATTTCCCGCGCAATGACG 1775
DB 472 CATCAGCAACACAGCAGATTCAGCAGTCGCGGAGAGAGATTTCCGTTGCAATTTGCG 531
QY 1776 AATAATGTTGGCTCTAAT 1835
DB 532 AATAACCATAGCTCTAAT 591
QY 1836 TTTTACAGTTTGAACGACAAATAGAAAATAGTTTAAAGATCTTTTATATATGCTTGT 1895
DB 592 TTTTTCAGTTTGAACGACAACTTAGAAAATAGTTTAAAGATCTTTTATATATGCTTGT 640
QY 1896 TGTGCTGCTGGTGAACAGTGTGATCTTTGATATATATATATATATATATATATATATAT 1955
DB 641 TGTGCTGCTGGTGAACAGTGTGATCTTTGATATATATATATATATATATATATATATAT 693
QY 1956 CTGTGTTAATTTCTTAAAGACTATTTT---TTAGTTTCCATTTGATGATAAATTTT 2008
DB 694 TTGGGTAATTTCTTATAAATTTTTCATAGTTTCGATTTTGGATAAATTTT 747

RESULT 8

ACC00806
 ID ACC00806 standard; cDNA; 1990 BP.
 XX
 AC ACC00806;
 XX
 DT 16-MAY-2003 (first entry)
 XX
 DE Glycine max oil trait related cDNA sequence SEQ ID NO:361.
 XX
 KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW LIP5-like transcription factor caleosin; ARP citrate lyase; SNF1;
 KW CKC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant; gene; ss.
 XX
 OS Glycine max.
 XX
 PN WO2003002751-A2.
 XX
 PD 09-JAN-2003.
 XX
 PF 27-JUN-2002; 2002WO-US020152.
 XX
 PR 29-JUN-2001; 2001US-0301913P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;
 XX
 XX WPI; 2003-201509/19.
 DR P-PSDB; ABR40771.
 XX
 PT Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX
 PS Claim 18; Page 381-382; 542pp; English.
 XX
 CC The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, LIP5-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity and CKC-like transcription factor activity. Also described: (1)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX
 XX Sequence 1990 BP; 628 A; 391 C; 459 G; 512 T; 0 U; 0 Other;
 SQ
 Query Match 19.2%; Score 385.8; DB 7; Length 1990;
 Best Local Similarity 76.9%; Pred. No. 2.1e-88;
 Matches 484; Conservative 0; Mismatches 142; Indels 3; Gaps 1;
 XX
 QY 692 TAGCGTTGACCTACCGAGAAACTATTGAGAGTTTGGACAGAGGACGCTATATA 751
 |||
 DB 560 TACTGTTGAGTTGCCTAGAGAACTTTGGATACATTGGGAGAGAACTCCATATA 619
 |||
 QY 752 CCGCGGTGTTACAGGATCGGTGACAGGAGATATGAGGCACATTATGGGATAATAG 811
 |||
 DB 620 TCGTGGAGTAACGACATAGATGGAAGGATGAACTCATCTTTGGGATATAG 679
 |||

812 TTCTAAACAGAGAGCGCCAAACCGCAAGAGAGACAGTTTATTGGGAGGTTATGACAA 871
 |||
 DB 680 CTGTAGAGGGAAGGCCAATCATGAAGAAAGGAGCCCAAGTTTATTGGGTGATATGATAA 739
 |||
 QY 872 AGAAGAAAAAGCAGCTAGGGCTTATGATTAGCCGCACTCAAGTATTGGGAACACCAC 931
 |||
 DB 740 AGAAGAGAAAGCAGCTAGAGCTTATGATTAGCTGCTGAAGTACTGGGGACATCCAC 799
 |||
 QY 932 TACTTACTTCTCCCATGAGCGAATATGAAAAGAGGTAGAGAGTGAAGCATGAC 991
 |||
 DB 800 CACTACCACTTTCCAAATTAGCACTATGAGAAGGAATTGGATGAATGAACACATGAC 859
 |||
 QY 992 AAGGCAAGAGTAGTTGCTCTCACTGCGCAGGAAAAAGTAGTGTCTCTCTCGTGTGATC 1051
 |||
 DB 860 GAGACAAGAAATTTGTTGCGCCATTAGAAGGAAGAGAGTGTCTCTCCAGGGGTGCATC 919
 |||
 QY 1052 GATTTATCGTGGAGTAACAAGACATCACCAATGGAAGATGGCAAGCTAGATAGGAAG 1111
 |||
 DB 920 AATGTATCGTGGAGTTACAAGGCATCACCAACACGGAAGATGGCAAGCAAGATTGGCAG 979
 |||
 QY 1112 AGTCGCGGTAAACAAAGACCTCTACTTGGGAACCTTTTGGCACACAAAGAGAGCTGCAGA 1171
 |||
 DB 980 AGTTGCAGGAACAAGATCTTTACTTGGGAACCTTTCACTAGTGAAGAGGCTGCAGA 1039
 |||
 QY 1172 GGCATACGACATTTGGGCGCCATCAAAATTCAGAGGATTAACCGAGTGAATACTTCGACAT 1231
 |||
 DB 1040 AGCATACGACATAGCAGCGATAAAGTTTCAGAGGTCTCAACGCTGTCACAAACTTTGACAT 1099
 |||
 QY 1232 GAACAGATACAACGTTAAAGCAATCTCTGAAAGCCCTAGTCTTCTCTAT---TGGTAGGCGC 1288
 |||
 DB 1100 GAGCGCTACGACGTGAAGGCCATTTCTTGAAGGACACACTCTCCCAATAGGAGGAGGCGC 1159
 |||
 QY 1289 CGCAAAACGTTCTCAAGGAGGCTAACCGTC 1317
 |||
 DB 1160 TCGAAAGGCTCTGAAAGAGGCTCAAGCTC 1188
 |||

RESULT 9
 ADD30878
 ID ADD30878 standard; cDNA; 1818 BP.
 XX
 AC ADD30878;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Plant yield-related polynucleotide clone G1793.
 XX
 KW ds; transcription factor; transgenic plant; growth rate; senescence;
 KW seed germination rate; plant vigor; seedling vigor.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO2003013227-A2.
 XX
 PD 20-FEB-2003.
 XX
 PF 09-AUG-2002; 2002WO-US025805.
 XX
 PR 09-AUG-2001; 2001US-0310847P.
 PR 19-NOV-2001; 2001US-0336049P.
 PR 11-DEC-2001; 2001US-0338692P.
 PR 14-JUN-2002; 2002US-00171468.
 XX
 PA (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX
 PI Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE.
 PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
 PI Brown PE;
 XX
 XX WPI; 2003-248221/24.
 DR P-PSDB; ADD30879.
 XX
 PT New plant transcription factor polynucleotides and polypeptides, useful

PT in producing transgenic plants with commercially valuable properties,
PT such as an alteration in a plant growth characteristic, e.g. growth rate
PT or apomixis.
XX
PS Disclosure; SEQ ID NO 907; 454pp; English.
XX
CC The invention relates to a number of isolated Arabidopsis thaliana cDNA
CC sequences and their encoded proteins which are especially transcription
CC factor related cDNA's and proteins. The isolated or recombinant plant
CC transcription factor polynucleotides and polypeptides are useful in
CC producing transgenic plants with commercially valuable properties, i.e.
CC modified or altered desirable traits as compared to a reference plant,
CC such as an alteration in a plant growth characteristic, e.g. growth rate,
CC germination rate of seeds, vigor of plants and seedlings, or leaf and
CC flower senescence. Sequence information related to the polynucleotides
CC and polypeptides can also be used in bioinformatic search methods. The
CC transgenic plant is useful for growing a progeny plant from a parent
CC plant. This sequence represents one of the cDNAs of the invention.
XX
SQ Sequence 1818 BP; 551 A; 425 C; 414 G; 428 T; 0 U; 0 Other;
Query Match 18.7%; Score 377.4; DB 9; Length 1818;
Best Local Similarity 76.7%; Pred. No. 2.8e-86;
Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;
QY 702 GCTACACCGAAGAACTATTGAGAGTTTGGACAGGACGCTCTATATACCGCGTGT 761
Db 554 GCCACGCCAAGACGTGCATTGGACACTTCGGACAACGAACTCGATCTATCGTGTGTC 613
QY 762 ACAAGGCATCGGTGGACAGGAAGATATGAGCACTTATGGGATATATGTTGTAAGA 821
Db 614 ACAAGCATCGGTGGACAGGAAGATATGAGCACTTATGGGATATATGTTGTAAGA 821
QY 822 GAAGGCCAAGCGGCAAGAACTATTGAGAGTTTGGACAGGACGCTCTATATACCGCGTGT 761
Db 614 ACAAGCATCGGTGGACAGGAAGATATGAGCACTTATGGGATATATGTTGTAAGA 821
QY 1002 TATGTTCCCTCACTGCGCAGGAAGAGTGTGTTCTCTCGTGTGTCATCGATTATCGT 1061
Db 854 TTGTTGCTGCTCCATTAGAGGAAAGTAGTGGATTTTCGAGAGCGCTTCGATGTCGA 913
QY 1062 GGAGTAAACAAGACATCAACCACTGGAAGATGGCAAGTGGCAAGGATCGGCCGCGGA 973
Db 914 GGAGTTACAAGGCATCAACCACTGGAAGATGGCAAGGATCGGCCGCGGA 973
QY 1122 AACAAAGACCTCTACTTGGGAACTTTGGACACAGAAAGCTCGAGGCTACGAC 1181
Db 974 AACAAAGACCTCTACTTGGGAACTTTGGACACAGAAAGCTCGAGGCTACGAC 1033
QY 1182 ATTCCGGCCATCAAAATTCAGAGGATTAACCGCAGTACTAATTCGACATCAACAGATAC 1241
Db 1034 ATAGCTGCAATAAGTTTAGAGGCTTAATGCACTGACCACTTCGAGATCAACCGGTAC 1093
QY 1242 AACGTTAAAGCAATTCCTGAAAGCCCTAGTCTTCCAT- TGTAGCGCGCAAAAGCT 1298
Db 1094 GACGTGAAAGCCATTCCTAGAGAGTAGCACTCTTCCCATCGAGGAGCGCAGCTAAACGG 1153
QY 1299 CTCAGGAGGCTAACCGTTC 1317
Db 1154 CTCAGGAGGCTCAAGCTC 1172
RESULT 10
ADD30822
ID ADD30822 standard; cDNA, 1818 BP.

XX
AC ADD30822;
XX
DT 15-JAN-2004 (first entry)
XX
DE Plant yield-related polynucleotide clone G1793.
XX
KW ds; transcription factor; transgenic plant; growth rate; senescence;
KW seed germination rate; plant vigor; seedling vigor.
XX
OS Arabidopsis thaliana.
XX
FN WO2003013227-A2.
XX
PD 20-FEB-2003.
XX
PF 09-AUG-2002; 2002WO-US025805.
XX
PR 09-AUG-2001; 2001US-0310847P.
PR 19-NOV-2001; 2001US-0336049P.
PR 11-DEC-2001; 2001US-0338692P.
PR 14-JUN-2002; 2002US-00171468.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
PI Brown FE;
XX
DR WPI; 2003-248221/24.
DR P-PSDB; ADD30823.
XX
PT New plant transcription factor polynucleotides and polypeptides, useful
PT in producing transgenic plants with commercially valuable properties,
PT such as an alteration in a plant growth characteristic, e.g. growth rate
PT or apomixis.
XX
PS Disclosure; SEQ ID NO 851; 454pp; English.
XX
CC The invention relates to a number of isolated Arabidopsis thaliana cDNA
CC sequences and their encoded proteins which are especially transcription
CC factor related cDNA's and proteins. The isolated or recombinant plant
CC transcription factor polynucleotides and polypeptides are useful in
CC producing transgenic plants with commercially valuable properties, i.e.
CC modified or altered desirable traits as compared to a reference plant,
CC such as an alteration in a plant growth characteristic, e.g. growth rate,
CC germination rate of seeds, vigor of plants and seedlings, or leaf and
CC flower senescence. Sequence information related to the polynucleotides
CC and polypeptides can also be used in bioinformatic search methods. The
CC transgenic plant is useful for growing a progeny plant from a parent
CC plant. This sequence represents one of the cDNAs of the invention.
XX
SQ Sequence 1818 BP; 551 A; 425 C; 414 G; 428 T; 0 U; 0 Other;
Query Match 18.7%; Score 377.4; DB 9; Length 1818;
Best Local Similarity 76.7%; Pred. No. 2.8e-86;
Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;
QY 702 GCTACACCGAAGAACTATTGAGAGTTTGGACAGGACGCTCTATATACCGCGTGT 761
Db 554 GCCACGCCAAGACGTGCATTGGACACTTCGGACAACGAACTCGATCTATCGTGTGTC 613
QY 762 ACAAGGCATCGGTGGACAGGAAGATATGAGCACTTATGGGATATATGTTGTAAGA 821
Db 614 ACAAGCATCGGTGGACAGGAAGATATGAGCACTTATGGGATATATGTTGTAAGA 821
QY 822 GAAGGCCAAGCGGCAAGAACTATTGAGAGTTTGGACAGGACGCTCTATATACCGCGTGT 761
Db 614 ACAAGCATCGGTGGACAGGAAGATATGAGCACTTATGGGATATATGTTGTAAGA 821
QY 882 GAAGGCCAAGCGGCAAGAACTATTGAGAGTTTGGACAGGACGCTCTATATACCGCGTGT 761
Db 674 GAAGGCCAAGCGGCAAGAACTATTGAGAGTTTGGACAGGACGCTCTATATACCGCGTGT 761
QY 882 GAAGGCCAAGCGGCAAGAACTATTGAGAGTTTGGACAGGACGCTCTATATACCGCGTGT 761
Db 734 GCAGCAAGATCATATGATCTAGCTAGCTAAAGTACTGGGTCCTCACTACTACTAAT 793

QY 942 TTCCCATGAGCGAATATGAAAGAGGTAGAGAGATGACCAATGACATGACAGGCGAG 1001
 DB 794 TTCCCATTAACAATGAGAGAAAGTAGAGAAATGAGCAATGACGAGCAAGAG 853
 QY 1002 TAGTTGCTCTACTCGCAGGAAAGTAGTGTCTCTCTGTTGTCATCGATTATCGT 1061
 DB 854 TTCTGTGGCTGCATTAGAGGAAAGTAGTGTATTTTCGAGAGGCGCTTCGATGATCGA 913
 QY 1062 GGAGTAAAGCAATCACCACATGAGATGCGAAGTACGATAGTAGAGAGTGCCTG 1121
 DB 914 GGAGTTACAGGCATCACCAATGAGATGCGAAGCAAGATGCGCGGAGTGCCTGCGGA 973
 QY 1122 AACAAAGACCTCTACTTGGGAATTTTGGCACAAAGAAAGCTGCGAGGCGCATACGAC 1181
 DB 974 AACAAAGACCTCTACTTGGGAATTTTGGCACTGAGGAAAGAGCAAGAGCTTACGAT 1033
 QY 1182 ATTGGGGCATCAAAATTCAGAGGATTAACCGAGTGAATTAATCTTCGATGATAC 1241
 DB 1034 ATAGCTGCAATAAAGTTTAGAGGACTTAATGAGTGAACCACTTCGAGATCAACCGGTAC 1093
 QY 1242 AACGTTAAAGCAATCTCGAAAGCCTAGTCTTCTAT---TGGTAGCGCGCAAAAGCT 1298
 DB 1094 GACGTGAAGCAATCTTAGAGGACTTCTTCCATCGAGAGGCGGAGCTTAAACGG 1153
 QY 1299 CTCAGGAGGCTAACCGTC 1317.
 DB 1154 CTCAAAGAGCTCAAGCTC 1172

RESULT 11

ID ADD30333 standard; cDNA; 1818 BP.
 XX

AC ADD30333;
 XX

DT 15-JAN-2004 (first entry)
 XX

DE Plant yield-related polynucleotide clone G1793.
 XX

KW ds; transcription factor; transgenic plant; growth rate; senescence;
 XX seed germination rate; plant vigor; seedling vigor.
 XX

OS Arabidopsis thaliana.
 XX

PN WO2003013227-A2.
 XX

PD 20-FEB-2003.
 XX

PF 09-AUG-2002; 2002WO-US025805.
 XX

PR 09-AUG-2001; 2001US-0310847P.
 XX

PR 19-NOV-2001; 2001US-0336049P.
 XX

PR 11-DEC-2001; 2001US-0338692P.
 XX

PR 14-JUN-2002; 2002US-00171468.
 XX

FA (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX

PI Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
 XX Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
 XX Broun PE;
 XX

DR WPI; 2003-248221/24.
 XX

DR P-PSDB; ADD30334.
 XX

PT New plant transcription factor polynucleotides and polypeptides, useful
 XX in producing transgenic plants with commercially valuable properties,
 XX such as an alteration in a plant growth characteristic, e.g. growth rate
 XX or apomixis.
 XX

PS Disclosure; SEQ ID NO 362; 454pp; English.
 XX

CC The invention relates to a number of isolated Arabidopsis thaliana cDNA

CC sequences and their encoded proteins which are especially transcription
 CC factor related cDNA's and proteins. The isolated or recombinant plant
 CC transcription factor polynucleotides and polypeptides are useful in
 CC producing transgenic plants with commercially valuable properties, i.e.
 CC modified or altered desirable traits as compared to a reference plant,
 CC such as an alteration in a plant growth characteristic, e.g. growth rate,
 CC germination rate of seeds, vigor of plants and seedlings, or leaf and
 CC flower senescence. Sequence information related to the polynucleotides
 CC and polypeptides can also be used in bioinformatic search methods. The
 CC transgenic plant is useful for growing a progeny plant from a parent
 CC plant. This sequence represents one of the cDNAs of the invention.
 XX

SQ Sequence 1818 BP; 551 A; 425 C; 414 G; 428 T; 0 U; 0 Other;

Query Match 18.7%; Score 377.4; DB 9; Length 1818;
 Best Local Similarity 76.7%; Pred. No. 2.8e-86;
 Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

QY 702 GTTACACCGAAGAAACTATTGAGAGTTTGGACAGAGACGCTCTATATACCGCGTGT 761
 DB 554 GCCAGCCCAAGACGTGCTTGCACACTTTCGACAAACGACCTCGATCTATCGTGTGTC 613
 QY 762 ACAAGCATCGTGGACAGAGATATGAGGACATTTATGGGTAATAGTTGTAAGA 821
 DB 614 ACAAGACATCGATGAGCTGTCGATATGAGGCTCATCTATGGGTAATAGTTGTAAGA 673
 QY 822 GAAGGCCAAACGCGCAAGAGAACAAAGTTTATTTGGGAGGTTATGACAAAGAGAAAAA 881
 DB 674 GAAGCCAGTCTAGGAAGAGAGCAAGTTTACTTTGGTGGATATGACAAAGAGATAA 733
 QY 882 GCAGTAGGGCTTATGATTAGCCGCACTCAAGTATTTGGGAAACCACTACTACTTAAC 941
 DB 734 GCAGCAAGATCATATGATCTAGCTGCTTAACTAGTACTGGGTCTTCACTACTACTAAT 793
 QY 942 TTCCCATGAGCGAATATGAAAAGAGGTAGAGAGATGAAGACATGACAAAGCAAGAG 1001
 DB 794 TTCCCATTTACAATCTACGAGAAAGTAGAGAAATGAAGACATGACGAGCAAGAG 853
 QY 1002 TAGTTGCTCTACTCGCAGGAAAGTAGTGTCTCTCTGTTGTCATCGATTATCGT 1061
 DB 854 TTCTGTGGCTGCCATTAGAAGGAAAGTAGTGTGATTTTCGAGAGGCGCTTCGATGTATCGA 913
 QY 1062 GGAGTAAAGCAATCACCACATGGAAGATGGCAAGCTAGATAGGAAGTGCCTG 1121
 DB 914 GGAGTTCAAGGCATCACCAATGGAAGATGGCAAGAGATGCGCCGAGTGCCTG 973
 QY 1122 AACAAAGACCTCTACTTGGGAACCTTTTGGCACAAAGAAAGCTGCGAGGCGCATACGAC 1181
 DB 974 AACAAAGACCTCTACTTGGGAACCTTTTAGCACTGAGGAAGAGCAGCAAGAGCTTACGAT 1033
 QY 1182 ATTGGCGGCATCAATTCAGAGGATTAACCGCAGTACTACTTTCGATCAAGAGATAC 1241
 DB 1034 ATAGCTGCAATAAAGTTTAGAGGACTTAATGAGTGAACAACTTCGAGATCAACCGGTAC 1093
 QY 1242 AACGTTAAAGCAATCTCGAAAGCCTAGTCTTCTCTAT---TGGTAGCGCGCAAAAGCT 1298
 DB 1094 GACGTGAAGCCATCTTAGAGATGAGCACTCTTCCCATCGGAGGAGCGGAGCTTAAACGG 1153
 QY 1299 CTCAGGAGGCTAACCGTC 1317
 DB 1154 CTCAAAGAGCTCAAGCTC 1172

RESULT 12

ACC00848

ID ACC00848 standard; cDNA; 1935 BP.

XX ACC00848;

XX ACC00848;

DT 16-MAY-2003 (first entry)

DE Gossypium hirsutum oil trait related cDNA sequence SEQ ID NO:486.

KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW LIP5-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CKC-like transcription factor; antisense inhibition; co-suppression;
 XX transgenic plant; gene; ss.
 OS Gossypium hirsutum.
 XX WO2003002751-A2.
 XX 09-JAN-2003.
 XX 27-JUN-2002; 2002WO-US020152.
 XX 29-JUN-2001; 2001US-0301913P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 XX Jones TU, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 XX Tarczynski MC;
 DR WPI; 2003-201509/19.
 DR F-PSDB; ABR40854.
 XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX Claim 18; Page 503; 542pp; English.
 XX The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, LIP5-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity, and CKC-like transcription factor activity. Also described: (I)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX Sequence 1935 BP; 576 A; 436 C; 416 G; 507 T; 0 U; 0 Other;
 SQ
 Query Match 17.7%; Score 357.4; DB 7; Length 1935;
 Best Local Similarity 73.8%; Pred. No. 48-81;
 Matches 454; Conservative 0; Mismatches 161; Indels 0; Gaps 0;
 QY 658 ACAACGTTGTGCGCAAGGCAAGCACTATTGATGATAGCGTTGAAGCTACACCGAAGAAA 717
 DB 442 AATCGATATTCANAAACACGACACACAGCCTATAGCTAGCGCTGCACCCCAAGAT 501
 QY 718 CTATTGAGATTGTCACAGAGAGCTCTATATACCGCGGTGTACAGGCATCGGTGA 777
 DB 502 CTGTTGATCTTTTGGCCAACTACCTCGATCTACCGGGGTGTCACCGGCATAGATGA 561
 QY 778 CAGGAGATATGCGCACATTTATGGGATATAGTTCTAAAGAGAGGCGCAACCGCA 837
 DB 562 CTGGAAGATATGAGCTCACTTGTGGGACATAGTTGTCAGAGAGAGGCGCAAGTAGGA 621
 QY 838 AAGGAAGACAGTTTATTGGGAGGTTATGCAAGAGAAAAGCAGCTAGGCTTATG 897
 DB 622 AAGGAAGCAAGTTTATTGGGTGCTATGACAGGAGATAAAGCTGCAAGAGCTTATG 681
 QY 898 ATTTAGCGGCACTCAAGTATTGGGAGACCACTACTACTACTTCCCATGAGCGAAT 957

DB 682 ATCTTGGCGCTCTCAAGTACTGGGGTCCGACCACTACTACCACTTTCCGATTTCCAACT 741
 QY 958 ATGAAAGAGGTAGAGATGAAGATGAAGACATGACAAAGGCAAGAGATATGTTCCCTCACTGC 1017
 DB 742 ATGAGAGAGGCTCGAAGAGATGAGAAACATGACTAGGCAAGAGTTGTTGTTCTCTCC 801
 QY 1018 GCAGGAAAAGTAGTGGTTTCTCTCGTGGTGCATCGATTTATCGTGGAGTAACAGACATC 1077
 DB 802 GGAGGAAAAGTAGTGGATTTTCTAGGGGTGCTTCAATTTACAGAGGAGTGACAAAGGCATC 861
 QY 1078 ACCAACATGGAAGATGCAAGCTAGGATAGGAAGAGTGCGCCGTAAACAAAGAGCCTTACT 1137
 DB 862 ATCAACATGTTAGTGGCAAGCAAGATTTGGAAGAGTTGCGAGGCAACAAAGATCTCTATC 921
 QY 1138 TGGGAACCTTTTGGCACACAAAGAAAGACTCCAGAGCATACGACATTCGGCCATCAAT 1197
 DB 922 TTGCACATTTAGCACCCAAAGAAAGAGCTGAAGCCTATGATATTGCGACCAATCAAGT 981
 QY 1198 TCAGAGATTAAACCGAGTGAAGTAACTTCGACATGAACAGATCAACAGTTTAAAGCAATCC 1257
 DB 982 TTAGAGGCTTTAAACGCGGTTACAAATTTTGATATGAGCCCTAGCTAGTAAACAGCATG 1041
 QY 1258 TCGAAAGCCCTAGTC 1272
 DB 1042 CAACAGCAATCTTC 1056
 RESULT 13
 ACC00804
 ID ACC00804 standard; cDNA; 2168 BP.
 XX ACC00804;
 XX 16-MAY-2003 (first entry)
 XX Glycine max oil trait related cDNA sequence SEQ ID NO:357.
 XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW LIP5-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CKC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant; gene; ss.
 XX Glycine max.
 XX WO2003002751-A2.
 XX 09-JAN-2003.
 XX 27-JUN-2002; 2002WO-US020152.
 XX 29-JUN-2001; 2001US-0301913P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 XX Jones TU, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 XX Tarczynski MC;
 XX WPI; 2003-201509/19.
 XX P-PSDB; ABR40769.
 XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX Claim 18; Page 377; 542pp; English.
 XX The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, LIP5-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity, and CKC-like transcription factor activity. Also described: (I)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX Sequence 1935 BP; 576 A; 436 C; 416 G; 507 T; 0 U; 0 Other;
 SQ

CC activated protein (MAP)-kinase activity, Lipid-like transcription factor
 CC activity, caldesin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity, and CKC-like transcription factor activity. Also described: (I)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (III), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC inhibition or co-suppression in a plant such as corn, soybean, wheat, rice,
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 2168 BP; 622 A; 570 C; 432 G; 544 T; 0 U; 0 Other;

Query Match 17.5%; Score 352; DB 7; Length 2168;
 Best Local Similarity 75.2%; Pred. No. 1e-79;
 Matches 439; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
 QY 700 AAGCTACACCGAAGAAACTATGAGGTTTGGACAGAGACGCTCTATATACCGCGGTG 759
 Db 517 AACATCCCCCAAGAAACCTCGACACCTTCGGGCAACGACCTCATCTACCGCGCG 576
 QY 760 TTACAGCGCTCGGTGGACAGAAATATGAGGACATTTATGGATTAATAGTTGTAATA 819
 Db 577 TGACCCGACATAGTAGGACTGGAGATACGAAGCTCATCTATGAGCAATAGTTGCAGAA 636
 QY 820 GAGAGGCGCAACCGCAAGAGAGACAAAGTTTATTTGGGAGGTTATGACAAACAGAAA 879
 Db 637 GGGAGGTCAAGCAGGAGAGAGAGCAAGTTTACTTGGTGGTTATGACAGAGGATA 696
 QY 880 AAGCAGTAGGCTTATGATTTACCGCACTCAAGTATTTGGGAACACCACTACTACTA 939
 Db 697 AGGCAGCCAGAGCTATGATCTCGAGCTCTCAAGTACTGGGTCCCACTACCACTA 756
 QY 940 ACTTCCCATCAGCGAATATGAAAGAGGTGAGAGATGACGACATGACAGGCAAG 999
 Db 757 ACTTCTCTATTTCACTATGAGAGAACTGGAGGAGATGAAGACATGACTAGGCAAG 816
 QY 1000 AGTATGTTGCTCTACTCGCAGGAAAAGTAGTGTGTTTCTCTCGTGTGTCATCGATTATC 1059
 Db 817 AGTTGTTGTTCTCTCTCGTAGGAAGAGCAGTGTGTTCTCTAGAGGGCTCTATATACA 876
 QY 1060 GTGGAGTAACAAGACATCCACATCGAAGATGGCAGTAGTAGTAGAGAGTGCAGG 1119
 Db 877 GAGGAGTAACGAGACACCAACGATGCGCGATGGCGGAGAGTAAGGAGAGTTGCGG 936
 QY 1120 GTAAACAAGACCTCTACTTGGGAACCTTTGGGCACACAAGAAAGCTGACAGGCGATACG 1179
 Db 937 GAAACAAGACCTCTACTTGGGACCTTTCAGACCCCAAGAGAGCTGTGAGGCGCTATG 996
 QY 1180 ACATTGGCGGCATCAATTTAGAGGATTAACCGAGTACTAATCTTCGACATGAACAGAT 1239
 Db 997 ACATTGCTGTATCAATTCAGGGGATTAATACGATGTAACAACTTTGACATGAGTCGCT 1056
 QY 1240 ACAACGTTAAAGCAATCCTCGAAAGCCCTAGTCTTCTCTATTGCT 1283
 Db 1057 ACGAGGTGAGAGCATTTGCAATAGTACTTCTCTATTGCTGCT 1100

RESULT 14

ACC00805
 ID ACC00805 standard; cDNA; 1290 BP.

XX
 AC ACC00805;

XX
 DT 16-MAY-2003 (first entry)

XX
 DE Glycine max oil trait related cDNA sequence SEQ ID NO:359.

XX
 KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;

KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW Lipid-like transcription factor caldesin; ATP citrate lyase; SNF1;
 KW CKC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant; gene; ss.
 XX
 OS Glycine max.
 XX WO2003002751-A2.
 XX
 XX 09-JAN-2003.
 PD
 XX 27-JUN-2002; 2002WO-US020152.
 XX
 XX 29-JUN-2001; 2001US-0301913P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;
 XX
 DR WPI; 2003-201509/19.
 XX
 XX F-PSDB; ABR40770.
 XX
 PT Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caldesin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX
 PS Claim 18; Page 379-380; 542pp; English.
 XX
 CC The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, Lipid-like transcription factor
 CC activity, caldesin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity, and CKC-like transcription factor activity. Also described: (I)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (III), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 1290 BP; 414 A; 231 C; 321 G; 324 T; 0 U; 0 Other;

Query Match 17.3%; Score 347.8; DB 7; Length 1290;
 Best Local Similarity 73.8%; Pred. No. 9.7e-79;
 Matches 442; Conservative 0; Mismatches 157; Indels 0; Gaps 0;
 QY 699 GAAGTACACCGAAGAAACTATTTGAGAGTTTGGACAGAGGACGCTCTATATACCGCGGT 758
 Db 117 GAACAGGTTCTCTCGCAATCTATTGACACTTTGGGCAAGAGACATCTCATATCGTGGT 176
 QY 759 GTTACAGGCGATCGGTGGACAGGAAGATATGAGCACATTTATGGGATAATAGTTGAAA 818
 Db 177 GTTACAGGCGATAGATGAGTGAAGATATGAGGCCCATTTGGGATAATAGTTGCGA 236
 QY 819 AGAGAAGGCCAAACGCGCAAGAGAGCAAGTTTATTTGGGAGGTTATGACAGAGAGAA 878
 Db 237 AAGGAGGGCAACCAAGAGAGAGGAGCAAGTTTACCTTGTGTTATGATAGGAGAA 296
 QY 879 AAGCAGCTAGGCTTATGATTTAGCGGCACTCAAGTATTTGGGGAACCACTACTACT 938
 Db 297 AAGCAGCAAAAGGCTTACGACCTTAGCTGCACTCAAGTATTTGGGTCCAACTCATA 356
 QY 939 AACTTCCCATGAGCGAATATGAAGAGGTGAGAGAGTGAAGACATGACACATGACAGGCA 998

Db 357 AATTTCCCTTTAAGCACCTTATGAGAGGAAGAACTTGAAGAGATGAAGACATGACCAGGCAA 416
Qy 999 GAGTATGTTGCTCACTCGCGAGGAAAGTAGTGTCTCTCTGTTGTCATCGATTAT 1058
Db 417 GAATTTGTCGCAATTAAGAGGAAGAGAGAGTGAATTTTCAAGAGGGGCACTGTGTAT 476
Qy 1059 CGTGGAGTAACAAGACATCACCAACATCGAAGATGGCAAGCTAGGATAGGAAGTCTGCC 1118
Db 477 AGAGTGTGACAGACACCATCAACATGGAAGGTGGCAGGCCCGCAATAGGAAGGTTGCA 536
Qy 1119 GGTACAAAGACCTTACTTGGGAACCTTTTGGCACACAAGAGAGCTGAGAGGCATAC 1178
Db 537 GGAACAAGAGACTTCTATCTTGTGTACATTTAGCACACAAGAGAGAGCTGCTGAAGCATAT 596
Qy 1179 GACATTGGGGCCATCAATTTAGAGGATTAACCGAGTGAATTAACCTTCGACATGAACAGA 1238
Db 597 GATATAGACAGCAATCAAGTTAGAGGACAAAGTGTGTGACCAACTTTGACATAAGTAGG 656
Qy 1239 TACACAGTTAAAGCAATCTCGAAGAGCCGTAGTCTTCTTATTTGTTAGCGCCGCAAAAG 1297
Db 657 TATGATGTGAAGAGATCTGCTCAAGCTCCACTCTTATTGCGAGGAGATCTTTGCAAAAG 715

RESULT 15
ACC00850
ID ACC00850 standard; cDNA; 1954 BP.
XX
AC ACC00850;
XX
16-MAY-2003 (first entry)
XX
DE Glycine max oil trait related cDNA sequence SEQ ID NO:490.
XX
KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
KW receptor-like protein kinase; mitogen activated protein kinase; oil;
KW LIP5-like transcription factor caleosin; ATP citrate lyase; SNF1;
KW CKC-like transcription factor; antisense inhibition; co-suppression;
KW transgenic plant; gene; ss.
XX
OS Glycine max.
XX
XX
PN WO2003002751-A2.
XX
PD 09-JAN-2003.
XX
PF 27-JUN-2002; 2002WO-US020152.
XX
PR 29-JUN-2001; 2001US-0301913P.
PR (DUPO) DU PONT DE NEMOURS & CO E I.
PA (PION-) PIONEER HI-BRED INT INC.
PA
XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
PI Tarczynski MC;
XX
DR WPI; 2003-201509/19.
DR P-PSDB; ABR40856.
XX

Novel nucleotide fragment encoding polypeptides having receptor-like protein kinase activity, caleosin-like activity, useful for altering oil phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
Claim 18; Page 508-509; 542pp; English.
XX
XX The present invention describes an isolated nucleotide fragment (I) comprising a nucleic acid sequence (NS) chosen from a NS encoding a polypeptide (PP) having receptor-like protein kinase activity, mitogen activated protein (MAP)-kinase activity, LIP5-like transcription factor activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like activity and CKC-like transcription factor activity. Also described: (1) activity and CKC-like transcription factor activity. (ii) comprising (I) or complement (II) of (I); (2) a chimeric construct (iii) comprising (I) or (ii), operably linked to a regulatory sequence; (3) a plant (iv)

CC comprising (iii) in its genome; (4) seeds (V) obtained from (iv); and (5) oil obtained from (v). (i) or its part can be used in antisense CC inhibiting or co-suppression in a transformed plant. (iii) is useful for altering the oil phenotype in a plant such as corn, soybean, wheat, rice, canola, Brassica, sorghum, sunflower or coconut. (iii) is also useful for creating transgenic plants having altered lipid profiles. (i) can also be used as a hybridisation probe. ACC00850 to ACC00868 and ABR40591 to ABR40879 represent sequences used in the exemplification of the present CC invention
XX
SQ Sequence 1954 BP; 582 A; 379 C; 468 G; 521 T; 0 U; 4 Other;
Query Match 17.3%; Score 347.8; DB 7; Length 1954;
Best Local Similarity 73.8%; Pred. No. 1.2e-78;
Matches 442; Conservative 0; Mismatches 157; Indels 0; Gaps 0;
Qy 699 GAAGCTACACCGAAGAAACATTTAGAGATTTTGGACAGAGGACGCTTATACCGCGGT 758
Db 781 GAACCACTTCTCCGAAATCTATTGACATTTTGGGCAAGAACATCTCAATATCGTGT 840
Qy 759 GTTACAAGGATCGTGGACAGGAAGATATGAGGACATTTATGGGATATAGTTGAAA 818
Db 841 GTTACAAGGATAGTGGACTGGAAGATATGAGGCCATTTTGGGATATATAGTTGAGA 900
Qy 819 AGAAGGCCAAACCGCGCAAGAAAGTATTTATTTGGAGGTTATGCAAGAGAA 878
Db 901 AAGCAAGGCAACACAGGAAGGAGCAAGTTTACCTTGTGTTATGATAAGGAAGAA 960
Qy 879 AAGCAGCTAGGCTTATGATTTAGCCGCACTCAAGTATTTGGGACCACTACTACT 938
Db 961 AAGCAGCAAGGCTTACGACTTAGCTGCACTCAAGTATTTGGGTCACACATCACA 1020
Qy 939 AACTTCCCATGAGCGAATATGAAAAGAGTGAAGAGATGAAGCAGCATGCAAGGCAA 998
Db 1021 AATTTCCCTTTAAGCACTTATGAGAGAACTTGAAGAGATGAAGCAGCATGCAAGGCAA 1080
Qy 999 GAGTATGTTGCTTCACTCGCGAGGAAAGTAGTGTCTCTCTGTTGTCATCGATTAT 1058
Db 1081 GAATTTGTCGCAATTTAAGAAAGGAGAGAGTGGATTTTCAAGAGGGGCACTCTGTAT 1140
Qy 1059 CGTGGAGTACACAGACATCCACATGGAAGATGGCAAGCTAGATAGGAAGTCTGCC 1118
Db 1141 AGAGTGTGACAGACACCATCAACATGGAAGGTTGGCAGGCCCGCAATAGGAAGGTTGCA 1200
Qy 1119 GGTAAACAAAGACCTCTACTTTGGGAACTTTTGGCACACAAGAAAGCTGCGAGGCTATC 1178
Db 1201 GGAAACAAGACTTGTATCTTTGTACATTTAGCACACAGAAGAAGCTGCTGAAGCATAT 1260
Qy 1179 GACATTGGGGCCATCAATTCAGAGGATTAACCGCAGTGTACTACTTCGACATGAACAGA 1238
Db 1261 GATATAGCAGCAATCAAGTTTCAGAGGAACAAAGTCTGTGACCAACTTTGACATAGTAGG 1320
Qy 1239 TACACGTTAAAGCAATCTCTGAAAGGCCCTAGTCTTCTTATTTGGTAGCGCGCAAAAG 1297
Db 1321 TATGATGTGAAGAGATCTGCTCAAGCTCCACTCTTATTTGCGAGGAGATCTTTGCAAAAG 1379

Search completed: July 19, 2004, 06:39:44
Job time : 828 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2004, 00:07:47 ; Search time 7901 Seconds
(without alignments)
11048.332 Million cell updates/sec

Title: US-09-980-364-1
Perfect score: 2014
Sequence: 1 gtcacatctctcttttaag.....gttgataaatttcagact 2014

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vt.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_in.*
- 18: em_hum.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vt.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	100.0	2014	2014	6	AX058687 Sequence
2	100.0	2014	2014	8	AF317904 Brassica
3	92.8	1888.6	1992	8	AF317905 Brassica
4	92.8	1888.6	2011	6	AX058689 Sequence
5	726	36.0	4873	8	AX058681 Sequence
6	726	36.0	4873	8	AF317906 Brassica
7	504.8	25.1	33563	8	ATT10B6
8	502.6	25.0	5151	6	AX058692 Sequence
9	502.6	25.0	5151	6	AF317907 Arabidops
10	399.2	19.8	1749	6	AX048247 Sequence
11	339.8	16.9	2546	8	AY461432 Nicotiana
12	330.6	16.4	1699	8	AY117207 Arabidops
13	330.6	16.4	1905	8	ATU41339
14	330.6	16.4	1955	8	ATU44028
15	330.6	16.4	2056	8	AY080706 Arabidops
16	330.6	16.4	2148	6	BD274516 Methods f
17	330.6	16.4	2148	6	AR316367 Sequence
18	330.6	16.4	2148	6	AR427901 Sequence
19	330.6	16.4	2148	8	U40256 Arabidops
20	327.2	16.2	2323	6	AX555218 Sequence
21	324.8	16.1	1738	6	BD274518
22	324.8	16.1	1738	6	AR316369 Sequence
23	324.8	16.1	1738	6	AR427903 Sequence
24	323.8	16.1	2153	8	AK111891
25	317.8	15.8	1745	8	AK101959 Oryza sat
26	315	15.6	1926	6	AX555220 Sequence
27	315	15.5	2510	8	AK106306 Oryza sat
28	308.2	15.3	2344	6	AX555216 Sequence
29	299.8	14.9	1773	8	ZMWHCF1
30	299	14.8	2164	8	AY062179 Oryza sat
31	269.2	13.4	1660	8	AK109839 Oryza sat
32	255.6	12.7	2010	6	AX555223 Sequence
33	250.2	12.4	1069	8	AY133880 Arabidops
34	250.2	12.4	1237	8	AY045915 Arabidops
35	246.2	12.2	1440	8	AK109848
36	228.4	11.3	1572	8	AK111617 Oryza sat
37	214	10.6	1866	8	AK106769 Oryza sat
38	205.2	10.2	1910	8	AK112088
39	192.8	9.6	912	8	ATH580379
40	192.8	9.6	2045	8	AK117300 Arabidops
41	190.6	9.5	1289	8	AK112070 Oryza sat
42	161.2	8.0	1494	8	BT008926 Triticum
43	152.2	7.6	1603	8	AY088118 Arabidops
44	145	7.2	1847	8	AF325506 Pisum sat
45	142.8	7.1	1395	6	AX505843 Sequence

ALIGNMENTS

RESULT 1
AX058687
LOCUS AX058687 2014 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 1 from Patent WO0075330.
ACCESSION AX058687
VERSION AX058687.1 GI:12311028
KEYWORDS
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1
AUTHORS Boutiller K., Ouellet, T., Custers, J., Hattori, J., Miki, B. and van
lookeren Campagne, M.

TITLE Use of the bmx3 transcriptional activator to control plant
embryogenesis and regeneration processes
JOURNAL Patent: WO 00/5330-A 1 14-DEC-2000;
Plant Research International (NL); THE MINISTER OF AGRICULTURE
(CA)

FEATURES Location/Qualifiers
source 1..2014
/organism="Brassica napus"
/mol_type="unassigned DNA"
/db_xref="taxon:3708"

ORIGIN

Query Match 100.0%; Score 2014; DB 6; Length 2014;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTCATCTCTCTTTAAGACCAAAACCTTTCTCTCTCTCTCTCTCATGATGAACCTTA 60
DB 1 GTTCATCTCTCTTTAAGACCAAAACCTTTCTCTCTCTCTCTCTCATGATGAACCTTA 60

QY 61 ACTAAGTCTCTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTCTCTGAACCTCAA 120
DB 61 ACTAAGTCTCTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTCTCTGAACCTCAA 120

QY 121 TGAATATAACTGGTTAGGCTTTCTCTCTCTCTCTCTATGACAAATACCATCGTAAGG 180
DB 121 TGAATATAACTGGTTAGGCTTTCTCTCTCTCTCTCTATGACAAATACCATCGTAAGG 180

QY 181 AGCTCTACTCTTCCACCAACCAACCGTCTGATAGTGTCCGCGAGAGTACTGTACGATC 240
DB 181 AGCTCTACTCTTCCACCAACCAACCGTCTGATAGTGTCCGCGAGAGTACTGTACGATC 240

QY 241 CGACCGCTCGCTCCGATGAGTCTTCAGCAATCCAAACATCTCTCTCTCTCTCTCTCTG 300
DB 241 CGACCGCTCGCTCCGATGAGTCTTCAGCAATCCAAACATCTCTCTCTCTCTCTCTCTG 300

QY 301 TCGTCTCGATCTCTTCCACGAGACAAACATAGTCACTCCGAGATGGGACATCAATG 360
DB 301 TCGTCTCGATCTCTTCCACGAGACAAACATAGTCACTCCGAGATGGGACATCAATG 360

QY 361 GTTGTGATGCAATTAACATCCACAGATGAGCAGATGGACCAACGATTCGAGATTTCC 420
DB 361 GTTGTGATGCAATTAACATCCACAGATGAGCAGATGGACCAACGATTCGAGATTTCC 420

QY 421 TTGGCGCACCAACCGATTTTCAACACCAACGAAACGTTTGGAGATGGAGTGAAGTG 480
DB 421 TTGGCGCACCAACCGATTTTCAACACCAACGAAACGTTTGGAGATGGAGTGAAGTG 480

QY 481 GCTGTTATGGAGGAGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
DB 481 GCTGTTATGGAGGAGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540

QY 541 GGTGAGAAATCAACCGGTGGATATGTGTATATCAAGAAATGGCAATGGTCAAAAG 600
DB 541 GGTGAGAAATCAACCGGTGGATATGTGTATATCAAGAAATGGCAATGGTCAAAAG 600

QY 601 GCTGTCTCTCTCAATGACT 660
DB 601 GCTGTCTCTCTCAATGACT 660

QY 661 AGCTTGTCTCCCAAGGAGACTATTGATGATGAGCTTGAAGCTTACACCAAGAAACTA 720
DB 661 AGCTTGTCTCCCAAGGAGACTATTGATGATGAGCTTGAAGCTTACACCAAGAAACTA 720

QY 721 TTGAGAGTTTGGACAGAGGAGCTCTATATACCGCGGTGTGTGTGTGTGTGTGTGTGTGT 780
DB 721 TTGAGAGTTTGGACAGAGGAGCTCTATATACCGCGGTGTGTGTGTGTGTGTGTGTGTGT 780

QY 781 GAAGATATGAGGACATTTATGGATATAGTTGTATTAAGAGAGGCAACCGCGCAAG 840
DB 781 GAAGATATGAGGACATTTATGGATATAGTTGTATTAAGAGAGGCAACCGCGCAAG 840

QY 841 GAAGACAGTTTATTTGGAGGTTATGACAAAGAAAGAAAGAGCTTAGGCTTATGATT 900

DB 841 GAAGACAGTTTATTTGGAGGTTTGTGACAAAGAAAGAAAGAGCTTAGGCTTATGATT 900

QY 901 TAGCGCAGCTCAAGTATTTGGGAAACCACTACTACTAACTTCCCATGAGGCAATATG 960

DB 901 TAGCGCAGCTCAAGTATTTGGGAAACCACTACTACTAACTTCCCATGAGGCAATATG 960

QY 961 AAAAAGAGGTAGAAGATGAAGCACATGACAAAGGCAAGAGTATGTTGCTCACTCGCA 1020

DB 961 AAAAAGAGGTAGAAGATGAAGCACATGACAAAGGCAAGAGTATGTTGCTCACTCGCA 1020

QY 1021 GGAAAGTAGTGGTTTCTCTCGTGTGTCATCGATTTATCTGTGAGTAACAAGACATCAC 1080

DB 1021 GGAAAGTAGTGGTTTCTCTCGTGTGTCATCGATTTATCTGTGAGTAACAAGACATCAC 1080

QY 1081 AACATGGAAGATGCAAGCTTAGGATAGGAAGAGTCCGCGGTAAACAAGACCTTACTTGG 1140

DB 1081 AACATGGAAGATGCAAGCTTAGGATAGGAAGAGTCCGCGGTAAACAAGACCTTACTTGG 1140

QY 1141 GAATTTTGGCACACAAGAGAGCTCGAGAGGATAGACATGTCGGGCAATCAAAATTC 1200

DB 1141 GAATTTTGGCACACAAGAGAGCTCGAGAGGATAGACATGTCGGGCAATCAAAATTC 1200

QY 1201 GAGGATTAACCGCAGTCACTAACTTTCGACATGAACAGATACAACTTAAAGCAATCCTCG 1260

DB 1201 GAGGATTAACCGCAGTCACTAACTTTCGACATGAACAGATACAACTTAAAGCAATCCTCG 1260

QY 1261 AAAGCCCTAGTCTTCTTATTTGGTAGGCGCCCAAAACGTTCTAAGAGGCTACCGTCCGG 1320

DB 1261 AAAGCCCTAGTCTTCTTATTTGGTAGGCGCCCAAAACGTTCTAAGAGGCTACCGTCCGG 1320

QY 1321 TTCAAGTATCATGATCATGATCACTAACTTTCAGAGAGTGAAGATAGTGTAGCGGTT 1380

DB 1321 TTCCAGTATGATGATGATGATCACTAACTTTCAGAGAGTGAAGATAGTGTAGCGGTT 1380

QY 1381 GCGAAACGCTCGGTTCCAGATCATCAGGAGTGAATTTGAGCTTATTTGCACCAATC 1440

DB 1381 GCGAAACGCTCGGTTCCAGATCATCAGGAGTGAATTTGAGCTTATTTGCACCAATC 1440

QY 1441 AAGAGAGTCAATGCTTATTTTACAAATGAGAGAAACCTTGTCTTCGAGAGTGTAGG 1500

DB 1441 AAGAGAGTCAATGCTTATTTTACAAATGAGAGAAACCTTGTCTTCGAGAGTGTAGG 1500

QY 1501 CTGTTTCAACAAGAGAGATGATCAACCACTTCTTGAGCAACAGAGAGCTCATGA 1560

DB 1501 CTGTTTCAACAAGAGAGATGATCAACCACTTCTTGAGCAACAGAGAGCTCATGA 1560

QY 1561 CTAATATCGATCATCAAAAGTTCTGTTTCGATGATTTACGTTTCTGTAAGAAATGTTG 1620

DB 1561 CTAATATCGATCATCAAAAGTTCTGTTTCGATGATTTACGTTTCTGTAAGAAATGTTG 1620

QY 1621 TTGTTATGTTGTTATCAAGGATTTGACCGCGGTAACTGCGATGCTTACGCTGCTA 1680

DB 1621 TTGTTATGTTGTTATCAAGGATTTGACCGCGGTAACTGCGATGCTTACGCTGCTA 1680

QY 1681 GTGAGTTTGAATATAACGCAAGAAACCATTTACTTGTCTGACAGCAGCAGACCCAG 1740

DB 1681 GTGAGTTTGAATATAACGCAAGAAACCATTTACTTGTCTGACAGCAGCAGACCCAG 1740

QY 1741 AGTCCGAGGTGGAGATTTCCGCGCGCAATGACGAATATGTTGCTCTAATATGATT 1800

DB 1741 AGTCCGAGGTGGAGATTTCCGCGCGCAATGACGAATATGTTGCTCTAATATGATT 1800

QY 1801 ACCATGGGAGGTGGTGGAGAAATGCTCCAAACATTTACGTTTGGAAACGAATAGA 1860

DB 1801 ACCATGGGAGGTGGTGGAGAAATGCTCCAAACATTTACGTTTGGAAACGAATAGA 1860

QY 1861 AAAATAGTTAAGATCTTTAGTATATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920

DB 1861 AAAATAGTTAAGATCTTTAGTATATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920

QY 1921 TTTGATTATGTTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1980

QY 1201 GAGGATTAACCGCAGTGAATACTTCGACATGAACAGATACAACTTAAGCAATCTCG 1260
 Db 1201 GAGGATTAACCGCAGTGAATACTTCGACATGAACAGATACAACTTAAGCAATCTCG 1260
 QY 1261 AAAGCCCTAGTCTCTTATGTTAGCGCGCAAAAGCTCTCAAGGAGGTAACTCCGCG 1320
 Db 1261 AAAGCCCTAGTCTCTTATGTTAGCGCGCAAAAGCTCTCAAGGAGGTAACTCCGCG 1320
 QY 1321 TTCACAGTATGATGATCAGTAAATGATGATGATGATGATGATGATGATGATGATG 1380
 Db 1321 TTCACAGTATGATGATCAGTAAATGATGATGATGATGATGATGATGATGATGATG 1380
 QY 1381 GGCACAAACCTCGGTTGACATCATCAGGAGTATGATGATGATGATGATGATGATG 1440
 Db 1381 GGCACAAACCTCGGTTGACATCATCAGGAGTATGATGATGATGATGATGATGATG 1440
 QY 1441 AAGAGAGGTACAAATGGTTATTAATCAATGGAGGAACTTGTCTTCGAGAGTGTAGG 1500
 Db 1441 AAGAGAGGTACAAATGGTTATTAATCAATGGAGGAACTTGTCTTCGAGAGTGTAGG 1500
 QY 1501 CTTGTTTCAACAAGAGGATGATCAACACCATTTCTTGGAGCAACACGAGAGCCTCATG 1560
 Db 1501 CTTGTTTCAACAAGAGGATGATCAACACCATTTCTTGGAGCAACACGAGAGCCTCATG 1560
 QY 1561 CTAATATCATCATCAAGTCTCTTTCGGATGATTCGGTACTGTTTGTGGAATGTTG 1620
 Db 1561 CTAATATCATCATCAAGTCTCTTTCGGATGATTCGGTACTGTTTGTGGAATGTTG 1620
 QY 1621 TTGGTTATGTTGTTATCAAGATTTGAGCCCGGTTAACTGCGATGCTTACGCTGCTA 1680
 Db 1621 TTGGTTATGTTGTTATCAAGATTTGAGCCCGGTTAACTGCGATGCTTACGCTGCTA 1680
 QY 1681 GTGAGTTTGAATTAACCAAGAAACCATTAATTAATTTCTGCTCAGCAGCAGACCCAGC 1740
 Db 1681 GTGAGTTTGAATTAACCAAGAAACCATTAATTAATTTCTGCTCAGCAGCAGACCCAGC 1740
 QY 1741 AGTCGCCAGGTGGAGATTTTCCCGCGCAATGACGAATAATGTTGGCTCTAATATGATT 1800
 Db 1741 AGTCGCCAGGTGGAGATTTTCCCGCGCAATGACGAATAATGTTGGCTCTAATATGATT 1800
 QY 1801 ACCATGGGAGGTGGGAGAGTCTCCAAATTTACAGTTTACAGTTTGAACGACAAATAGA 1860
 Db 1801 ACCATGGGAGGTGGGAGAGTCTCCAAATTTACAGTTTACAGTTTGAACGACAAATAGA 1860
 QY 1861 AAAAATAGTTAAAGATCTTATGTTATATGCGTTGTTGTTGCTGGTGAACAGTGTATAC 1920
 Db 1861 AAAAATAGTTAAAGATCTTATGTTATATGCGTTGTTGTTGCTGGTGAACAGTGTATAC 1920
 QY 1921 TTTCATATGTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 1980
 Db 1921 TTTCATATGTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 1980
 QY 1981 TTTAGTTTCCATAGTTGGATAAATTTTCAGACT 2014
 Db 1981 TTTAGTTTCCATAGTTGGATAAATTTTCAGACT 2014

RESULT 3
 AF317905
 LOCUS 1992 bp mRNA linear PLN 10-OCT-2002
 DEFINITION Brassica napus AP2/EREBP transcription factor BABY BOOM2 (BBM2)
 mRNA, complete cds.
 ACCSSION AF317905
 VERSION AF317905.1 GI:21069052
 KEYWORDS
 SOURCE Brassica napus (rape)
 ORGANISM Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 (bases 1 to 1992)
 AUTHORS Boutilier, K., Offringa, R., Sharma, V.K., Kieft, H., van
 Lammeren, A.A.M., Zhang, L., Hattori, J., Ouellet, T., Liu, C.-M.,

Miki, B.L.A., Custers, J.B.M. and van Lookeren Campagne, M.M.
 Ecopic expression of the Brassica napus BABY BOOM gene triggers a
 conversion from vegetative to embryonic growth

Unpublished

2 (bases 1 to 1992)

Boutilier, K., Ouellet, T. and Hattori, J.

Direct Submission

Submitted (02-NOV-2000) Business Unit Plant Development and

Reproduction, Plant Research International, Droeveendaalsesteeg 1,

Wageningen 6708 PB, The Netherlands

Location/Qualifiers

1. 1992

/organism="Brassica napus"

/mol_type="mRNA"

/cultivar="Tropas"

/db_xref="taxon:3708"

1. 1992

/gene="BBM2"

51. 1790

/gene="BBM2"

/codon_start=1

/product="AP2/EREBP transcription factor BABY BOOM2"

/protein_id="AA033801.1"

/db_xref="GI:21069053"

/translation="MNNWLGFSLPYEQNHHRKDVCSSTTTTAVDAGECYDPTAA

SDSAIQTSPFPFVLDATFRDNNSHRDWDINGSACNNIHNDQDQKLENFLG

RTTITNTNENVDIDGGCGGSLGSLMIKTLRNQPDVNDQENKENGFLG

GLSLWNSSTSCDNNYSNNLVAOKTIDDSVEATPKTIESQFPTSIYGVTRHR

WTGRYEALWNSCKRGQTRGVYLGVDKEEKAARAYDLAAKHWGTTTTNFP

MSEYKEIEBKHMTQBYVAISLRKSSGFSRGASIRGVTRHHQHWQRIQIRG

NKDLKFTGTQEEAAEAYDAIAIKRGLTATVFNPMNVKAILLESPLIGSAK

RIKEANRPVPMNMISNNVSENNASQWNAVQHGGVDSLLQHQHRYNGYVQ

GNSSERSAFCRQEDDQHLFNTQSLMTNIHQSSVSDSVTCNVVGVGYGQ

FAAPVNCDAAYASBPFDYARNHYVFAQQQQTQHSFGDFPAAMTNVNSNMYHGE

GEVAPTFVWMDN"

ORIGIN

Query Match 92.8%; Score 1868.6; DB 8; Length 1992;
 Best Local Similarity 97.9%; Pred. No. 0;
 Matches 1904; Conservative 0; Mismatches 39; Indels 2; Gaps 1;
 QY 70 TTTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTCTGAACTCAATGAATA 129
 Db 1 TTTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTCTGAACTCAATGAATA 60
 QY 130 ACTGTTAGGCTTTCTCTCTCTTATGAACAAAATCACCATCGTAAGACGCTTACT 189
 Db 61 ACTGTTAGGCTTTCTCTCTCTTATGAACAAAATCACCATCGTAAGACGCTTACT 120
 QY 190 CTTCCACCACCAACCGTCTGATGTCGCGGAGAGTACTGTAGCATCCGACCGCTG 249
 Db 121 CTTCCACCACCAACCGCTGATGTCGCGGAGAGTACTGTAGCATCCGACCGCTG 180
 QY 250 CTTCCGATGAGTCTTCCAGCCATCCAAACATCGTTCTCTCTCTCTCTCTCTCT 309
 Db 181 CTTCCGATGAGTCTTCCAGCCATCCAAACATCGTTCTCTCTCTCTCTCTCTCT 240
 QY 310 ATGTTTCCAGAGACAAATAGTCACTCCGAGATTCGGACATCAATGGTGTGAT 369
 Db 241 ATGTTTCCAGAGACAAATAGTCACTCCGAGATTCGGACATCAATGGTGTGAT 300
 QY 370 GCAATAATCCCAACAGATGAGCAAGATGACCAAGCTTGAGAATTTCTTGGCGCA 429
 Db 301 GCAATAATCCCAACAGATGAGCAAGATGACCAAGCTTGAGAATTTCTTGGCGCA 360
 QY 430 CCACGAGATTTCAACCAACGAAACGTTGAGATGGAAGTGAAGTGGTGTATG 489
 Db 361 CCACGAGATTTCAACCAACGAAACGTTGAGATGGAAGTGAAGTGGTGTATG 420
 QY 490 GAGGAGAGCGGTGGTGGTGGTCTACTAGGACTTTTCGATGATAAAGACATGCTGAGAA 549
 Db 421 GAGGAGAGCGGTGGTGGTGGTCTACTAGGACTTTTCGATGATAAAGACATGCTGAGAA 480

QY 550 ATCAACCGGTGGATATGTTGATATCAAGAAAATGGCAATGCTGCAAAAGCGCTGTGCC 609
 Db 481 ATCAACCGGTGGATATGTTGATATCAAGAAAATGGCAATGCTGCAAAAGCGCTGTGCC 540
 QY 610 TCTCAATGAACCTCATCTACTCTTGTGATAACCAACGACAGCAATAAACAAGCTGTGTTG 669
 Db 541 TCTCAATGAACCTCATCTACTCTTGTGATAACCAACGACAGCAATAAACAAGCTGTGTTG 600
 QY 670 CCCAAGGGAAGACTATTTGATGATAGCGTTGAAGCTACACGGAAGAAAATCTATTGAGAGTT 729
 Db 601 CCCAAGGGAAGACTATTTGATGATAGCGTTGAAGCTACACGGAAGAAAATCTATTGAGAGTT 660
 QY 730 TTGGACAGAGAGACCTCTATATACCGCGGTCTTCAAGAGCATCGGTGGACAGGAAGATATG 789
 Db 661 TTGGACAGAGAGACCTCTATATACCGCGGTCTTCAAGAGCATCGGTGGACAGGAAGATATG 720
 QY 790 AGGCACATTTATGGGTAATAGTTGTAAGAGAGAGCGCCAAACGCGCAAGAGAGCAAG 849
 Db 721 AGGCACATTTATGGGTAATAGTTGTAAGAGAGAGCGCCAAACGCGCAAGAGAGCAAG 780
 QY 850 TTTATTTGGGAGGTTATGACAAAGAGAAAAGCAGCTAGGGCTTATGATTTAGCGGCAC 909
 Db 781 TTTATTTGGGAGGTTATGACAAAGAGAAAAGCAGCTAGGGCTTATGATTTAGCGGCAC 840
 QY 910 TCAAGTATTTGGGAAACCAACCTACTACTACTTCCCATGACGCAATATGAAAGAGG 969
 Db 841 TCAAGTATTTGGGAAACCAACCTACTACTACTTCCCATGACGCAATATGAAAGAGG 900
 QY 970 TAGAAGAGATGAAGCATGACAGGCAAGAGATGTTGCTCACTGCGCAGGAAAAGTA 1029
 Db 901 TAGAAGAGATGAAGCATGACAGGCAAGAGATGTTGCTCACTGCGCAGGAAAAGTA 960
 QY 1030 GTGGTTCTCTCGTGTGCAATCGATTTATCGTGAGTAAAGAGATACCAACATGGA 1089
 Db 961 GTGGTTCTCTCGTGTGCAATCGATTTATCGTGAGTAAAGAGATACCAACATGGA 1020
 QY 1090 GATGGCAAGCTAGGATAGGAGAGTCCCGGTAAAGAGCTCTACTTTGGGAACCTTTTG 1149
 Db 1021 GATGGCAAGCTAGGATAGGAGAGTCCCGGTAAAGAGCTCTACTTTGGGAACCTTTTG 1080
 QY 1150 GCACACAGAGAGAGCTGACAGAGCATACGACATGCGGCCATCAAAATCAGAGGATTA 1209
 Db 1081 GCACACAGAGAGAGCTGACAGAGCATACGACATGCGGCCATCAAAATCAGAGGATTA 1140
 QY 1210 CCGCAGTGACTAACTTCGACATGAACAGATACACGTTAAAGCAATCTCGAAAGCCCTA 1269
 Db 1141 CCGCAGTGACTAACTTCGACATGAACAGATACACGTTAAAGCAATCTCGAAAGCCCTA 1200
 QY 1270 GTCTTCTTATTTGGTAGCGCGCAAAACGCTCTCAAGAGGCTAACCGTCCGGTCCAGTA 1329
 Db 1201 GTCTTCTTATTTGGTAGCGCGCAAAACGCTCTCAAGAGGCTAACCGTCCGGTCCAGTA 1260
 QY 1330 TGATGATGATCAGTAATAACGTTTCAGAGAGTGAGATAGTGTAGCGGTTGGCAAAACG 1389
 Db 1261 TGATGATGATCAGTAATAACGTTTCAGAGAGTGAGATAGTGTAGCGGTTGGCAAAACG 1320
 QY 1390 CTGGGTTTACGATCATCAGGAGTAGATTTAGCTTATTTGACCAACATCAAGAGAGT 1449
 Db 1321 CTGGGTTTACGATCATCAGGAGTAGATTTAGCTTATTTGACCAACATCAAGAGAGT 1380
 QY 1450 ACAATGTTTATTTAACAATGGAGAACTTGTCTTCGAGAGTGCTAGGCGCTTTGTTC 1509
 Db 1381 ACAATGTTTATTTAACAATGGAGAACTTGTCTTCGAGAGTGCTAGGCGCTTTGTTC 1440
 QY 1510 AACAGAGGATGATCAACACATTTCTTGAGCAACACGAGAGGCTCATGACTAATATCG 1569
 Db 1441 AACAGAGGATGATCAACACATTTCTTGAGCAACACGAGAGGCTCATGACTAATATCG 1500
 QY 1570 ATCATCAAGTCTGTTTCGATGATTTCCGTTACTGTTTGTGGAATGTTGTTGTTATG 1629
 Db 1501 ATCATCAAGTCTGTTTCGATGATTTCCGTTACTGTTTGTGGAATGTTGTTGTTATG 1560
 QY 1630 GTGGTTATCAAGGATTTGACAGCCCGGTTAACTCCGAGTCCCTAGCTGTAGTGAAGTTG 1689

Db 1561 GTGGTTATCAAGGATTTGCAGCCCGGTTAACTGCGATGCCCTACGCTGTAGTGAAGTTG 1620
 QY 1690 ATTATAACGCAAGAAACCAATTTACTTTGCTCAGCAGCAGAGACCCACAGTCCGCCAG 1749
 Db 1621 ACTATAACGCAAGAAACCAATTTACTTTGCTCAGCAGCAGAGACCCACAGTCCGCCAG 1680
 QY 1750 GTGGAGATTTTCCCGGGCAATGACGAATAATGTTGGCTCTAATATGTTATTAACCATGGG 1809
 Db 1681 GAGGAGATTTTCCCGGGCAATGACGAATAATGTTGGCTCTAATATGTTATTAACCATGGG 1740
 QY 1810 AAGTGCTGGAGAGTTGCTCCAACTTTACAGTTTGGAAACGCAATTAAGAAAATAGT 1869
 Db 1741 AAGTGCTGGAGAGTTGCTCCAACTTTACAGTTTGGAAACGCAATTAAGAAAATAGT 1800
 QY 1870 TAAAGATCTTTAGTATATGCGTTGTTGCTGCTGGTGAACAGTGTGATCTTGAATAT 1929
 Db 1801 TAAAGATCTTTAGTATATGCGTTGTTGCTGCTGGTGAACAGTGTGATCTTGAATAT 1860
 QY 1930 GTTTTTTCTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1989
 Db 1861 GTTTTTT- TTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1918
 QY 1990 CATTAGTTGGATAAATTTTCAGACT 2014
 Db 1919 CATTAGTTGGATAAATTTTCGACT 1943
 RESULT 4
 LOCUS AX058689 2011 bp DNA linear PAT 17-JAN-2001
 DEFINITION Sequence 3 from Patent WO0075330.
 ACCESSION AX058689
 VERSION AX058689.1 GI:12311029
 KEYWORDS Brassica napus (rape)
 SOURCE
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE
 AUTHORS
 Boutlier,K., Ouellet,T., Custers,J., Hattori,J., Miki,B. and van
 Lookeren Campagne,M.
 TITLE
 Use of the bnm3 transcriptional activator to control plant
 embryogenesis and regeneration processes
 JOURNAL
 Patent: WO 0075330-A 3 14-DEC-2000;
 Plant Research International (NL) ; THE MINISTER OF AGRICULTURE
 (CA)
 FEATURES
 Location/Qualifiers
 1..2011
 /organism="Brassica napus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3708"
 ORIGIN
 Query Match 92.8%; Score 1868.6; DB 6; Length 2011;
 Besc Local Similarity 97.9%; Pred. No. 0;
 Matches 1904; Conservative 0; Mismatches 39; Indels 2; Gaps 1;
 QY 70 TTCTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTCTCTGAACTCAATGAATAATA 129
 Db 1 TTCTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTTTCTGAACTCGATGAATAATA 60
 QY 130 ACTGGTTAGCTTTCT 189
 Db 61 ACTGGTTAGCTTTTCT 120
 QY 190 CTTCCACCAACCAACCGCTGTAGATGTCCCGAGAGTACTGTTACGATCCGACCGCTG 249
 Db 121 CTTCCACCAACCAACCGCTGTAGATGTCCCGAGAGTACTGTTACGATCCGACCGCTG 180
 QY 250 CTTCCGAGAGTCTTCAGCCATCAACATCGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 309

Db 181 CCTCGATGAGTCTTCCAGCCATCCAAACATCGTTCTCTCTCCCTTGGTGTCTGCTCCG 240
QY 310 ATGCTTTTCCAGAGACAAATAGTCACTCCCGAGATTGGGACATCAATCGTTGTGCAT 369
Db 241 ATGCTTTTCCAGAGACAAATAGTCACTCCCGAGATTGGGACATCAATCGTTGTGCAT 300
QY 370 GCAATACATCCCAACAGATGAGAGATGAGACCAAAAGCTTGGAGATTTCTTGGCCGCA 429
Db 301 GTAATTAACATCCCAATAGTATGAGAGATGAGACCAAAAGCTTGGAGATTTCTTGGCCGCA 360
QY 430 CCACACAGATTTTACACACCAACGAAACGTTGGAGATGGAAGTGGAGTGGTGTATG 489
Db 361 CCACACAGATTTTACACACCAACGAAACGTTGGAGATTCGATGGAGTGGTGTATG 420
QY 490 GAGGAGAGACGGTGGTGTGCTCACTAGAGCTTTCGATGATTAAGACATGGCTGAGAA 549
Db 421 GAGGAGAGACGGTGGTGTGCTCACTAGAGCTTTCGATGATTAAGACATGGCTGAGAA 480
QY 550 ATCAACCCGTGGATATGTTGATATCAAGAAATGGCAATGCTGCAAAAGCCCTGTCCC 609
Db 481 ATCAACCCGTGGATATGTTGATATCAAGAAATGGCAATGCTGCAAAAGCCCTGTCCC 540
QY 610 TCTCAATGAATCTACTCTTCTGTGATTAACAAACGACAGCAATTAACACCTTTG 669
Db 541 TCTCAATGAATCTACTCTTCTGTGATTAACAAACGACAGCAATTAACACCTTTG 600
QY 670 CCCAAGGGAAGACTATTGATGATAGCGTTGAAGCTTACACCGAAGAAATATTTCAGAGTT 729
Db 601 CCCAAGGGAAGACTATTGATGATAGCGTTGAAGCTTACACCGAAGAAATATTTCAGAGTT 660
QY 730 TTGGACAGAGGAGCTCTATATACCGCGGTGTGTACAGGCTATCGGTGACAGAGATATG 789
Db 661 TTGGACAGAGGAGCTCTATATACCGCGGTGTGTACAGGCTATCGGTGACAGAGATATG 720
QY 790 AGGCACATTTATGGGATATAGTCTGTAAGAGAGAGGCCAAACCGCAAGAGAGACAAG 849
Db 721 AGGCACATTTATGGGATATAGTCTGTAAGAGAGAGGCCAAACCGCAAGAGAGACAAG 780
QY 850 TTTATTTGGAGGTATGACAAAGAGAGAGAGAGAGCTAGGCTTATGATTTAGCCGCAC 909
Db 781 TTTATTTGGAGGTATGACAAAGAGAGAGAGAGAGCTAGGCTTATGATTTAGCCGCAC 840
QY 910 TCAAGTATTTGGGAAACCACTACTACTACTCTTCCCATGAGCGAATATGAGAAAGAGG 969
Db 841 TCAAGTATTTGGGAAACCACTACTACTACTCTTCCCATGAGCGAATATGAGAAAGAGG 900
QY 970 TAGACAGATGAGACATGACAGGAGAGATGTTGCTCTACTGCGCAGGAAAGTA 1029
Db 901 TAGACAGATGAGACATGACAGGAGAGATGTTGCTCTACTTGGCAGGAAAGTA 960
QY 1030 GTGGTTTCTCTGTTGTGATCGATCGATTTATCGTGGAGTAAACAGACATCACCACATGAA 1089
Db 961 GTGGTTTCTCTGTTGTGATCGATCGATTTATCGTGGAGTAAACAGACATCACCACATGAA 1020
QY 1090 GATGCGAAGCTAGGATGAGAGAGTTCGCGGTAAACAGACCTCTACTTGGAACTTTG 1149
Db 1021 GATGCGAAGCTAGGATGAGAGAGTTCGCGGTAAACAGACCTCTACTTGGAACTTTG 1080
QY 1150 GCACACAAAGAGAGCTGACAGAGGATACGACATTTGCGGCCATCAAAATTCAGAGGATTA 1209
Db 1081 GCACACAAAGAGAGCTGACAGAGGATACGACATTTGCGGCCATCAAAATTCAGAGGATTA 1140
QY 1210 CGCAGTGTACTTTCGACATGACAGATCAAGCTTTAAAGCAATCTCGAAGCCCTA 1269
Db 1141 CGCAGTGTACTTTCGACATGACAGATCAAGCTTTAAAGCAATCTCGAAGCCCTA 1200
QY 1270 GTCTTCTTATTTGGTACGCGCGCAAAAGCTCTCAAGAGGCTTAACCGTCCGTTCCAAAGTA 1329
Db 1201 GTCTTCTTATTTGGTACGCGCGCAAAAGCTCTCAAGAGGCTTAACCGTCCGTTCCAAAGTA 1260
QY 1330 TGNATGATGATGATATACGTTTTCAGAGAGTGAAGATGCTAGCGGTTGGCAAAAG 1389
Db 1261 TGNATGATGATGATATACGTTTTCAGAGAGTGAAGATGCTAGCGGTTGGCAAAAG 1320

QY 1390 CTGCGGTTTCCAGCATCATCAGGAGTAGATTGAGCTTATTGCAACCATCAAGAGAGGT 1449
Db 1321 CTGCGGTTTCCAGCATCATCAGGAGTAGATTGAGCTTATTGCAACCATCAAGAGAGGT 1380
QY 1450 ACAATGGTTATATATACATGAGGAAACTTGTCTCGAGAGTGTAGGCTTGTTC 1509
Db 1381 ACAATGGTTATATATACATGAGGAAACTTGTCTCGAGAGTGTAGGCTTGTTC 1440
QY 1510 AACAGAGAGATGATCAACACCATTTCTTGAACCAACGACAGAGCCTCATGACTAATATCG 1569
Db 1441 AACAGAGAGATGATCAACACCATTTCTTGAACCAACGACAGAGCCTCATGACTAATATCG 1500
QY 1570 ATCATCAAGATCTGTTTTCGAGATGATTCGGTACTTGTGTTGAAATGTTGGTATG 1629
Db 1501 ATCATCAAGATCTGTTTTCGAGATGATTCGGTACTTGTGTTGAAATGTTGGTATG 1560
QY 1630 GTGGTTATCAAGGATTTTCCAGCCCGGTTAACTGCGATGCTACGCTGCTAGTGTGTTG 1689
Db 1561 GTGGTTATCAAGGATTTTCCAGCCCGGTTAACTGCGATGCTACGCTGCTAGTGTGTTG 1620
QY 1690 ATTATACGAGAGAAACCATTTACTTGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 1749
Db 1621 ACTATACGAGAGAAACCATTTACTTGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1750 GTGAGAGATTTTCCCGCGGCAATGACGAATATGTTGGCTCTAATATGATTTACCATGGG 1809
Db 1681 GAGGAGATTTTCCCGCGGCAATGACGAATATGTTGGCTCTAATATGATTTACCATGGG 1740
QY 1810 HAGTGTGGAGAGATGCTCCACATTTACGTTTGGAGAGAGAGAGAGAGAGAGAGAGAG 1869
Db 1741 HAGTGTGGAGAGATGCTCCACATTTACGTTTGGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 1870 TAAAGATCTTTAGTATATGCTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1929
Db 1801 TAAAGATCTTTAGTATATGCTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
QY 1930 GTTCTTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1989
Db 1861 GTTCTTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1918
QY 1990 CATTAGTTGGATAAATTTTCAGACT 2014
Db 1919 CATTAGTTGGATAAATTTTCAGACT 1943

RESULT 5
AX058691
LOCUS 4873 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 5 from Patent WO0075330.
ACCESSION AX058691
VERSION AX058691.1 GI:12311030
KEYWORDS Brassica napus (rape)
SOURCE Brassica napus
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1
AUTHORS Boutilier, K., Ouellet, T., Custers, J., Hattori, J., Miki, B. and van
lookeren Campagne, M.
TITLE Use of the bms3 transcriptional activator to control plant
JOURNAL embryogenesis and regeneration processes
Patent: WO 0075330-A 5 14-DEC-2000;
Plant Research international (NL); THE MINISTER OF AGRICULTURE
(CA)
FEATURES
source Location/Qualifiers
1 .4873
/organism="Brassica napus"
/mol_type="unassigned DNA"
/db_xref="taxon:3708"
1620 ..1622
/note="start codon"
misc_feature

intron 1846..2298
intron 2720..2952
intron 3036..3160
intron 3170..3214
intron 3404..3553
intron 3628..3797
intron 3849..3961
intron 4039..4148
misc_feature 4856..4858
/note="stop codon"
ORIGIN
Query Match 36.0%; Score 726; DB 6; Length 4873;
Best Local Similarity 100.0%; Pred. No. 1.3e-162;
Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1149 GGCACACAGAGAGAGCTGAGAGGATACGACATTCGCGCCATCAAAATTCAGAGGATTA 1208
Db 4148 GGCACACAGAGAGAGCTGAGAGGATACGACATTCGCGCCATCAAAATTCAGAGGATTA 4207
QY 1209 ACCGAGTGAATTAATTCGACATGAACAGATACACGTTAAAGCAATCTCGAAGCCCT 1268
Db 4208 ACCGAGTGAATTAATTCGACATGAACAGATACACGTTAAAGCAATCTCGAAGCCCT 4267
QY 1269 AGTCCTTCATTTGGTAGCGCCGACAAACGCTCTCAAGAGGCTAACCGTCCGGTTCACAGT 1328
Db 4268 AGTCCTTCATTTGGTAGCGCCGACAAACGCTCTCAAGAGGCTAACCGTCCGGTTCACAGT 4327
QY 1329 ATGATGATGATCATGATTAATTAACGTTTCAGAGAGTGAGATAGTCTAGCGGTTGGCAAAAC 1388
Db 4328 ATGATGATGATCATGATTAATTAACGTTTCAGAGAGTGAGATAGTCTAGCGGTTGGCAAAAC 4387
QY 1389 GCTGCGGTTGAGTATCATCAGGAGTAGATTGAGTTATGAGTTATGACCAACATCAAGAGAGG 1448
Db 4388 GCTGCGGTTGAGTATCATCAGGAGTAGATTGAGTTATGAGTTATGACCAACATCAAGAGAGG 4447
QY 1449 TACAATGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1508
Db 4448 TACAATGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4507
QY 1509 AAACAGAGGATGATCAACACCAATTTCTTGAGCAACACGAGAGCTCATGACTAATATC 1568
Db 4508 AAACAGAGGATGATCAACACCAATTTCTTGAGCAACACGAGAGCTCATGACTAATATC 4567
QY 1569 GATCATCAAGATGCTGTTTCGGATGATTCGGTACTGTTTGTGGAATGTTGTGTTAT 1628
Db 4568 GATCATCAAGATGCTGTTTCGGATGATTCGGTACTGTTTGTGGAATGTTGTGTTAT 4627
QY 1629 GGTGGTTATCAAGATTTGAGCCCGGTTAACTGGATGCTTACGCTGCTAGTGATTT 1688
Db 4628 GGTGGTTATCAAGATTTGAGCCCGGTTAACTGGATGCTTACGCTGCTAGTGATTT 4687
QY 1689 GATTATACGCAAGAAACCAATTTACTGTTGCTCAGCAGCAGACCCAGCAGTCCCA 1748
Db 4688 GATTATACGCAAGAAACCAATTTACTGTTGCTCAGCAGCAGACCCAGCAGTCCCA 4747
QY 1749 GGTGGATTTTCCCGCGCAATGACGAATTAATGTTGGCTCTAATATGTTATACCATGG 1808
Db 4748 GGTGGATTTTCCCGCGCAATGACGAATTAATGTTGGCTCTAATATGTTATACCATGG 4807
QY 1809 GAAGTGGTGAGAGTTGCTCCACATTTACAGTTTGGACCAATTAAGAAATAG 1868
Db 4808 GAAGTGGTGAGAGTTGCTCCACATTTACAGTTTGGACCAATTAAGAAATAG 4867
QY 1869 TTAAG 1874
Db 4868 TTAAG 4873

RESULT 6
AF317906
LOCUS
DEFINITION Brassica napus AP2/EREBP transcription factor BABY BOOM1 (BBM1)
linear
4873 bp
DNA
PLN 10-OCT-2002

gene, complete cds.
AF317906
AF317906.1 GI:21069054
KEYWORDS
SOURCE
ORGANISM
Brassica napus (rape)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Eucalyptaceae;
1 (bases 1 to 4873)
Boutillier, K., Offringa, R., Sharma, V. K., Kieft, H., van
Lammeren, A. A. M., Ouellet, T., Zhang, L., Hattori, J., Liu, C.-M.,
Miki, B. L. A., Custers, J. B. M., and van Lookeren Campagne, M. M.
Ecopic expression of BABY BOOM triggers a conversion from
vegetative to embryonic growth
Plant Cell 14 (8), 1737-1749 (2002)
22163009
12172019
2 (bases 1 to 4873)
Boutillier, K. and Fiers, M.
Direct Submission
Submitted (02-NOV-2000) Business Unit Plant Development and
Reproduction, Plant Research International, Droevendaalsesteeg 1,
Wageningen 6708 PB, The Netherlands
Location/Qualifiers
1. 4873
/organism="Brassica napus"
/mol_type="genomic DNA"
/cultivar="Topas"
/db_xref="taxon:3708"
1620..4858
/gene="BBM1"
join(1620..1845,2299..2719,2953..3035,3161..3169,
3315..3403,3554..3627,3798..3848,3962..4038,4149..4858)
/gene="BBM1"
/product="AP2/EREBP transcription factor BABY BOOM1"
join(1620..1845,2299..2719,2953..3035,3161..3169,
3315..3403,3554..3627,3798..3848,3962..4038,4149..4858)
/gene="BBM1"
/codon_start=1
/product="AP2/EREBP transcription factor BABY BOOM1"
protein_id="AA33802.1"
/db_xref="GI:21069055"
/translation="MNNWLGSLSPYEQNHRRKDVYSGSTTTTVDVAGEVCYDPTAA
SDSASATQSPFPFVVVADFTDNNSHSDWDINGCANNNHNDQDQPKLENFLG
RTTYTINENVGDSGSGSGGGGSLGLSMITLNRQVNDVNDGNGNAK
GLSLMSNSTSDNNNDNNVAVQKTIIDSVATPKTIESFQRTSIYRGVTRHR
WTGVEAHLWNSCKREGQTRGVYLGVDKEEKAARYDLAALKYWGTTTTFP
MSBEKEVEEMKMTROYVAISLREKSGFSGASIVRGVTRHGHQWQARIQVAG
NKDLYLFTFGQEEAERAYDIAAIKFRGLIATVTPDMNRVNVKAILSPSPISGAAK
RLKANRPVPSMMIISNNVSENSASGQNAVAQHFGQVDLSLHQRQERYNGYYN
CGNLSSEARACFKQEDDQHFHLSMTNIDHQSVDSDSVTCGNVVGYYGQ
PAAPVNCDAVAASFPDYNARNHYVFAQQQTQSPGDFPAAMTNNVGSNNMYHGE
GEVAPTFVTWNDN"
ORIGIN
Query Match 36.0%; Score 726; DB 8; Length 4873;
Best Local Similarity 100.0%; Pred. No. 1.3e-162;
Matches 726; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1149 GGCACACAGAGAGAGCTGAGAGGATACGACATTCGCGCCATCAAAATTCAGAGGATTA 1208
Db 4148 GGCACACAGAGAGAGCTGAGAGGATACGACATTCGCGCCATCAAAATTCAGAGGATTA 4207
QY 1209 ACCGAGTGAATTAATTCGACATGAACAGATACACGTTAAAGCAATCTCGAAGCCCT 1268
Db 4208 ACCGAGTGAATTAATTCGACATGAACAGATACACGTTAAAGCAATCTCGAAGCCCT 4267
QY 1269 AGTCCTTCATTTGGTAGCGCCGACAAACGCTCTCAAGAGGCTAACCGTCCGGTTCACAGT 1328
Db 4268 AGTCCTTCATTTGGTAGCGCCGACAAACGCTCTCAAGAGGCTAACCGTCCGGTTCACAGT 4327
QY 1329 ATGATGATGATCAGTAATAACGTTTCAGAGAGTGAGATAGTCTAGCGGTTGGCAAAAC 1388

```

Db      4328 ATGATGATGATCAGTAATAACGTTTCAGAGAGTGAGAAATAGTGTAGCGGTGGCAAAAC 4387
QY      1389 GTGCGGTTACGATCATCAGGAGTAGATTGAGCTTTATTGCACCAACATCAAGAGAGG 1448
Db      4388 GCTGCGGTTACGATCATCAGGAGTAGATTGAGCTTTATTGCACCAACATCAAGAGAGG 4447
QY      1449 TACATGTTATTATTAAGTGGGAACTGTCTTCGGAGAGTCTAGGCTGTGTTTC 1508
Db      4448 TACATGTTATTATTAAGTGGGAACTGTCTTCGGAGAGTCTAGGCTGTGTTTC 4507
QY      1509 AAACAAGAGGATGATCAACACCATTTCTTGAGCAACAGCCAGAGCCTCATGACTAATATC 1568
Db      4508 AAACAAGAGGATGATCAACACCATTTCTTGAGCAACAGCCAGAGCCTCATGACTAATATC 4567
QY      1569 GATCATCAAGTCTGTTTCGGATGATTCGGTTACTGTTTGTGAAATGTTGTTGTTAT 1628
Db      4568 GATCATCAAGTCTGTTTCGGATGATTCGGTTACTGTTTGTGAAATGTTGTTGTTAT 4627
QY      1629 CGTGTGTTATCAAGGATTTGCGAGCCCGGTTAACTGCGATGCTACGCTAGTGTAGTGT 1688
Db      4628 CGTGTGTTATCAAGGATTTGCGAGCCCGGTTAACTGCGATGCTACGCTAGTGTAGTGT 4687
QY      1689 GATTATACGCAAGAAACATTTACTTGTCTGAGCAGCAGCAGCCAGCAGTCGCCA 1748
Db      4688 GATTATACGCAAGAAACATTTACTTGTCTGAGCAGCAGCAGCCAGCAGTCGCCA 4747
QY      1749 GGTGAGAGATTTCCCGCGGCAATGACGAATATGTTGCTCTAATATGTTATACCATGGG 1808
Db      4748 GGTGAGATTTCCCGCGGCAATGACGAATATGTTGCTCTAATATGTTATACCATGGG 4807
QY      1809 GAAGTGTGAGAGATTTGCTCCAAATTTACGTTTGGACGACAAATTAGAAAAATAG 1868
Db      4808 GAAGTGTGAGAGATTTGCTCCAAATTTACGTTTGGACGACAAATTAGAAAAATAG 4867
QY      1869 TTAAG 1874
Db      4868 TTAAG 4873

RESULT 7
LOCUS   ATT10B6/c
DEFINITION Arabidopsis thaliana DNA chromosome 5, BAC clone T10B6 (ESSA project).
ACCESSION AL391142
VERSION   2
KEYWORDS  Arabidopsis thaliana (thale cress)
SOURCE   Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 33563)
Sato, S., Nakamura, Y., Kaneko, T., Kato, T., Asamizu, E., Kotani, H.,
Tabata, S., Mewes, H. W., Rudd, S., Lemcke, K. and Mayer, K. F. X.
Unpublished
2 (bases 1 to 33563)
EU Arabidopsis Sequencing project.
Direct Submission
Submitted (07-AUG-2000) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lencke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de
Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/.
location/Qualifiers
1. .33563
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="5"

```

```

gene
CDS
2657. .3208
/gene="T10B6_10"
2657. .3208
/gene="T10B6_10"
/note="similarity to unknown protein T17B22.3 - Arabidopsis thaliana, EMBL:AC012328"
/codon_start=1
/product="putative protein"
/protein_id="CAC01730.1"
/db_xref="GI:9755758"
/db_xref="SPTREMBL:Q9LF49"
/translation="MGNVSSGALSKTSSSSSSAAKVI LPDGGVRNIHAPMKAELMME IPSFLVDAKSLIKRKFCEPLAADDLOIKGCHVVAFPMTSATSAANASLARI.FVA AKQRBHVSGSDHSAAYKHCHNNGRVSPDGGDDVKVISAGSKINLEIDIEFSAAEF MARI SVSKSKPKLEIIVVEESLS"
2657. .3208
/gene="T10B6_10"
/number=1
4424. .4951
/gene="T10B6_20"
4424. .4951
/gene="T10B6_20"
/codon_start=1
/product="putative protein"
/protein_id="CAC01731.1"
/db_xref="GI:9755759"
/db_xref="SPTREMBL:Q9LF48"
/translation="MGDOSI SNKLPDEKXVMDLTCGOEALILEELRICLPKWVTOEQVT DMIRGTCNIVEIDNVFYEHTKYEQVSAVISTFRDGISSSNHDAFVSKLEIFHE TERSSCQPSQFHLKSSSSQKSSISPVNKRKI KGPKKKKASSKVESGGPKQAS ITKFFNKVSSSEGTKS"
4424. .4951
/gene="T10B6_20"
/number=1
5605. .8523
/gene="T10B6_30"
complement(join(5605. .5681,5778. .5888,6000. .6117, 6198. .6255,6342. .6432,6606. .6672,6861. .6961,7118. .7220, 7369. .7575,7734. .7864,7949. .8092,8233. .8357,8457. .8523))
/gene="T10B6_30"
complement(join(5605. .5681,5778. .5888,6000. .6117, 6198. .6255,6342. .6432,6606. .6672,6861. .6961,7118. .7220, 7369. .7575,7734. .7864,7949. .8092,8233. .8357,8457. .8523))
/gene="T10B6_30"
/note="Contains Prokaryotic membrane lipid attachment site AA144-154"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAC01732.1"
/db_xref="GI:9755760"
/db_xref="SPTREMBL:Q9LF47"
/translation="MVSDEKHEQSRQVHNKMPKLPFGYDEKKNRYFPKGPPIP GAKSSSRITQKPEKPEQETNYKQTKLKLXVLSRELNGNIVSNKKNMFRDE IQKTQAYPVVRYGSTEIDGTALQKQVDPVGLTRKXILVAGSAGCGLILRV SKQKQVGVIECDPVSLPKCKENDTEERAPHLRPTCLVALSSISIELIGS DASNSKPVNLTFRSLQGVNSECITWSDCNISGSHAAIGTDIAGLVLETVG SYFLRSKSDVPALQFHSGNIVHCGLANGALVSDLRERPAFLTSHNILEFQLQGNIN PSHVYFSSLTCLTKTSIQYLMASSMDGLDQRMVKRGVGVQTYEGHVNSTPI EFGIDPSERFILSGDDCYTRIVSIKSQLSENKFNSPVSVVVCWSADERQDRKDS IGHGAWLGSREAFNLL"
complement(5605. .5681)
/gene="T10B6_30"
/number=1
complement(5682. .5777)
/gene="T10B6_30"
/number=1
complement(5778. .5888)
/gene="T10B6_30"
/number=2
complement(5889. .5999)
/gene="T10B6_30"
/number=2
complement(6000. .6117)

```


[illegible]

Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Arabidopsis Open Reading Frame (ORF) Clones

TITLE JOURNAL REFERENCE AUTHORS

2 (bases 1 to 1699)
Yamada, K., Ban, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (03-JUN-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

TITLE JOURNAL

COMMENT
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-Length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Yamada, K., Ban, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to Genbank.

FEATURES source

Location/Qualifiers
1..1699
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosomes="4"
/clone="U11873"
/ecotype="Columbia"
/notes="This clone is in pUNI 51."
1..1699
/gene="At4g37750"
1..1668
/gene="At4g37750"
/codon_start=1
/evidence=experimental
/product="putative ovule development protein aintegumenta"
/protein_id="AA051282.1"
/db_xref="GI:21436287"
/translation="MKSFDNDNHNSTNLLGFLSSNMKMGGRGREALYSSST
SRTATSSVPVPLVGDNTSFGVCGYSGNPNNGGIYSHSVPLSDSLCMEALNR
SSHNHDSRPKVEDPFGTHNNTSHKAMDLSLDFYNTHPNTTNTQEFESP
PQTRHEETRNYPDPSLTHGGSFNVGVYGFQSLSLSPGSSQSCITGSHHQ
NQNQSQNHQOQISEALVETSVGFETTTMAAKKRGQEDVYVVGQKQIVHRKSIDT
FGQRTQYGVTRHRTGYEAHLWDNSFKRGHRSKQVLYGGVDEEKAARAYDL
AALKYWGSPHTFSAENYQKETEDMKNTROYVAHLRKRSGSFGSRGASIVRGYTRH
HQHGRQARTRVAGNKDLYLGFQCEAAEAYDAVAIKFRGTNAVTFDITRYDVID
RIMSNSTLSELARLNNSIVVRNEDQTLNAVVEGSGNKEVSPPELLSFPALFA
LPQVNMFGSNGMGNPWTSPNPAELKVALTLFQMPVFAANADS"
1669..1699
/gene="At4g37750"

gene CDS

1905 bp mRNA linear PLN 23-OCT-1996
Arabidopsis thaliana ANT (AINTEGUMENTA) mRNA, complete cds.
ATU41339
Arabidopsis thaliana ANT (AINTEGUMENTA) mRNA, complete cds.
U41339
U41339.1 GI:1244707
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1905)
Elliot, R.C., Betzner, A.S., Huttner, E., Oakes, M.P., Tucker, W.Q., Gerentes, D., Perez, P. and Smyth, D.R.
AINTEGUMENTA, an APTALA2-like gene of Arabidopsis with pleiotropic roles in ovule development and floral organ growth
Plant Cell 8 (2), 155-168 (1996)
96351414
MEDLINE
8742707
PUBMED
2 (bases 1 to 1905)
Smyth, D.R.
Direct Submission
Submitted (27-NOV-1995) David Smyth, Genetics and Dev. Biology, Monash University, Wellington Road, Clayton, VIC 3168, Australia
Location/Qualifiers
1..1905
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Landsberg erecta"
/db_xref="taxon:3702"

ORIGIN

Query Match 16.4%; Score 330.6; DB 8; Length 1699;
Best Local Similarity 75.4%; Pred. No. 4.9e-68;
Matches 411; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 714 AAAAATATTGAGAGTTTGGACAGAGAGAGCTCTATATACCGCGGTGTTTCAAGAGCATCGG 773
Db 814 AAATCTATCGATCTTTTGGACAAAGAACTTCTCATACCGAGCGGTTTCAAGACATAGA 873
QY 774 TGGACAGGAGATATGAGGCACATTTTATGGGATATAGTTTAAAGAGAGAGAGAGAGG 833
Db 874 TGGACTGGTAGATATGAAGCTCATCTATGGGACAAATAGTTTCAAGAGAGAGAGGTCACAGT 933
QY 834 CGCAAGGAGAGACAAAGTTTATTTGGGAGGTTATGACAAAGAGAGAGAGAGAGAGAGGCT 893
Db 934 AGAAAGGAGAGACAAAGTTTATCTGGGAGGTTATGATATGGAGAGAGAGAGAGAGAGGCT 993
QY 894 TATGATTTATCCCGACCTCAAGTATTTGGGAGAACCAACCACTACTACTACTTCCCATGAGC 953
Db 994 TATGATCTTCTGCACTCAAGTACTGGGTCCTCTACTCACCACCAATTTCTCTGGGAG 1053
QY 954 GAATATCAAAAAGAGGTAGAGAGATGAACACATGACAGGAGAGAGAGAGAGAGTATGTCCTCA 1013
Db 1054 AATTTATCAGAAAGAGATGAAGACATGAAGACATGACTAGACAGATATGTTGCACAT 1113
QY 1014 CTGGCAGGAGAAAGTAGTGTGTTTCTCTCGTGGTCATCGATTTTATCGTGGAGTAACAAGA 1073
Db 1114 TTGAGAGGAGAGAGAGAGTGTGTTTCTCTAGGGTGTCTTCCATCTATAGAGGAGTACACAAGA 1173
QY 1074 CATCACCAACATGGAAGATGCGAGCTAGCATAGGAGAGAGTCCCGGTACAAAGACCTC 1133
Db 1174 CATCACAGCATGGAAGGTGGCAAGCAGGATTTGGTAGATCGCTGGAAACAAAGATCTC 1233
QY 1134 TACTTTGGGAACCTTTTGGCACACAAAGAAAGAGCTGCAGAGGCATACGACATTTGGGGCCATC 1193
Db 1234 TACCTTGGAACTTTTGGAAACCCAGAGAGAGTGCAGAGAGCTTACGATGTAGCAGCAATT 1293
QY 1194 AAATTCAGAGATTAAACCGAGTACTACTTGGACATGACATGACATGACATGACATGACATGACAT 1253
Db 1294 AAGTTCCGTGGCACAAATGCTGTGACTAACTTTGATATCAGAGGTACGATGTTGATCGT 1353
QY 1254 ATCCT 1258
Db 1354 ATCAT 1358

RESULT 13

ATU41339
LOCUS
DEFINITION
Arabidopsis thaliana ANT (AINTEGUMENTA) mRNA, complete cds.
ACCESSION
U41339
VERSION
U41339.1
KEYWORDS
GI:1244707
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana

REFERENCE AUTHORS

ELLIOTT, R.C., BETZNER, A.S., HUTTNER, E., OAKES, M.P., TUCKER, W.Q., GERENTES, D., PEREZ, P. and SMYTH, D.R.
AINTEGUMENTA, an APTALA2-like gene of Arabidopsis with pleiotropic roles in ovule development and floral organ growth
Plant Cell 8 (2), 155-168 (1996)
96351414
MEDLINE
8742707
PUBMED
2 (bases 1 to 1905)
Smyth, D.R.
Direct Submission
Submitted (27-NOV-1995) David Smyth, Genetics and Dev. Biology, Monash University, Wellington Road, Clayton, VIC 3168, Australia
Location/Qualifiers
1..1905
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Landsberg erecta"
/db_xref="taxon:3702"

FEATURES source

Location/Qualifiers
1..1905
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Landsberg erecta"
/db_xref="taxon:3702"

```

/chromosome="4"
/map="between spt and cer-9"
/tissue type="flower; ovule"
1..1905
/gene="AINTEGUMENTA"
38..1705
/gene="AINTEGUMENTA"
/note="similar to APETALA2 protein encoded by GenBank
Accession Number U12546"
/codon_start=1
/product="ANT"
/protein_id="AAB17364.1"
/db_xref="GI:1244708"
/translation="MKSFCDNDNHNNTNLLGFLSSNMVMKGRGGREAIYSST
SSAATSSSVPPQLVVDGNTSNFGVCYGNPNNGIYSHMVMPLRSDGSLCLMEALNR
SSHNHHDQSDSPKVEDPFGTHNNTSHKEAMDLSLDFYNTTHEPNTTNFOEFTSF
POTNHEETNYGNDPSLTHGGSFNVGVGEFOQSLSLMSPGSSSCITGSHHHQ
NONHONSONHQOISEALVETSVGFETTAAAKKRGQEDVVVGKQIVHRKSIDT
FGQTSQYGVTRHWTRGYEAHLWDSFKKEGSRKRGROYLGGIDMEKAAKADL
AALKWGPSTHNFSAENYKQIEIDMKNTQREYVAHLRKRKSSGSRGASIEYGVTRH
HQRGWARIGRVAGNKDLYLGTGTEQEEAAEDVAAIKFRGTNAVTFDITRYDVD
RIMSNTLLSGELARRNNSIVVRNTEDQTLNNAVVEGSKNEVSTPERLLSPFAIFA
LPQVNMFGSMGNNMSPNTSNPNAELKTVALTLTPMPVFAAWADS"
878..1108
/misc_feature
/gene="AINTEGUMENTA"
/note="encodes first AP2 domain"
1184..1391
/misc_feature
/gene="AINTEGUMENTA"
/note="encodes second AP2 domain"

ORIGIN
Query Match 16.4%; Score 330.6; DB 8; Length 1905;
Best Local Similarity 75.4%; Pred. No. 4.9e-68;
Matches 411; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 714 AAAAATAATTGAGAGTTTGGACAGAGCGTCTATATACCGCGGTGTACAAAGGCATCGG 773
DB 851 AAATCTATCGATCTTTTGGACACGAGCTTCTCAATACCGAGCGTTCACAGACATAGA 910
QY 774 TGGACAGAGATATAGGACATTTATGGATAATAGTTGTAAGAGAGAGCCAAACG 833
DB 911 TGGACTGATATAGAGTCACTATGATGGACAATAGTTTCAAGAGGAGGTCACAGT 970
QY 834 CGCAAAGGAAGACAAGTTTATTTGGAGGTTATGACAAAGAAAGAAAGCAGTACGGCT 893
DB 971 AGAAAAGGAAGACAAGTTTATCTGGAGGTTATGATATGGAGGAGAGCTGTCAGCA 1030
QY 894 TATGATTTAGCGGACTCAAGTATTTGGGAAACCACTACTACTACTACTACTACTACTACT 953
DB 1031 TATGATCTTCTGCACTCAAGTACTGGGTCCCTCTACTCACCACCAATTTCTCTGGGAG 1090
QY 954 GAATATGAAAAGAGGTAGAGAGATGAAGACATGACAAAGGCAAGAGTATGTCCTCA 1013
DB 1091 AATTATCAGAAAGAGATGAAGACATGAAGAAATGACTAGACAGAAATATGTCACAT 1150
QY 1014 CTGGCAGGAAAAGTACTGTTTCTCTCGGGTGCATGCAATTTATCGTGAGTAAACAAG 1073
DB 1151 TTGAGAGGAGGAGCAGTGGTTTCTCTAGGGGTGCTTCCATCTATAGAGAGTCAACA 1210
QY 1074 CATCACCAATGGAAGTCGCAAGTATGATAGGAGAGTCCCGGTAAACAAAGACCTC 1133
DB 1211 CATCACCAATGGAAGTGGCAGACGAGATTTGATAGTCTGCTGAAAACAAAGATCTC 1270
QY 1134 TACTTTGGAACTTTTGGCAACAAGAGAGTTCGAGAGGCATACGACATTCGGCCATC 1193
DB 1271 TACCTTTGGAACTTTTGGAAACCAAGAGAGTTCGAGAGCTTACGATGTAGCAGCAAT 1330
QY 1194 AAATTACAGAGATTAACCGAGTGACTTACTTCGACATGACAGATACAGATTAAGCA 1253
DB 1331 AGTTTCGTGGCAACAATGCTGTGACTTACTTTGATATCAGGAGTACGATGTTGATCGT 1390
QY 1254 ATCCT 1258
|||||

```

```

Db 1391 ATCAT 1395

RESULT 14
ATU44028
LOCUS
DEFINITION
Arabidopsis thaliana transcription factor CKC mRNA, complete cds.
ACCESSION
U44028
VERSION
U44028.1 GI:1171428
KEYWORDS
SOURCE
Arabidopsis thaliana (Chale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1955)
REFERENCE
Vergani, P., Morandini, P. and Soave, C.
Direct Submission
Submitted (29-DEC-1995) Paola Vergani, Dept. of Biology 'L.
Gorini', University of Milan, Via Celoria 26, Milan, 20133, Italy
Location/Qualifiers
1..1955
/organism="Arabidopsis thaliana"
/mo_type="mRNA"
/db_xref="taxon:3702"
76..1743
CDS
/function="transcription factor; complements a yeast
protein kinase C 1 mutant; Method: conceptual translation
supplied by author."
/codon_start=1
/product="CKC"
/protein_id="AAA86281.1"
/db_xref="GI:1171429"
/translation="MKSFCDNDNHNNTNLLGFLSSNMVMKGRGGREAIYSST
CSAATSSSVPPQLVVDGNTSNFGVCYGNPNNGIYSHMVMPLRSDGSLCLMEALNR
SSHNHHDQSDSPKVEDPFGTHNNTSHKEAMDLSLDFYNTTHEPNTTNFOEFTSF
POTNHEETNYGNDPSLTHGGSFNVGVGEFOQSLSLMSPGSSSCITGSHHHQ
NONHONSONHQOISEALVETSVGFETTAAAKKRGQEDVVVGKQIVHRKSIDT
FGQTSQYGVTRHWTRGYEAHLWDSFKKEGSRKRGROYLGGIDMEKAAKADL
AALKWGPSTHNFSAENYKQIEIDMKNTQREYVAHLRKRKSSGSRGASIEYGVTRH
HQRGWARIGRVAGNKDLYLGTGTEQEEAAEDVAAIKFRGTNAVTFDITRYDVD
RIMSNTLLSGELARRNNSIVVRNTEDQTLNNAVVEGSKNEVSTPERLLSPFAIFA
LPQVNMFGSMGNNMSPNTSNPNAELKTVALTLTPMPVFAAWADS"

ORIGIN
Query Match 16.4%; Score 330.6; DB 8; Length 1955;
Best Local Similarity 75.4%; Pred. No. 4.9e-68;
Matches 411; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 714 AAAAATAATTGAGAGTTTGGACAGAGCGTCTATATACCGCGGTGTACAAAGGCATCGG 773
DB 889 AAATCTATCGATCTTTTGGACACGAGCTTCTCAATACCGAGCGTTCACAGACATAGA 948
QY 774 TGGACAGAGATATAGGACATTTATGGATAATAGTTGTAAGAGAGAGCCAAACG 833
DB 949 TGGACTGATATAGAGTCACTATGAGGACAATAGTTTCAAGAGGAGGTCACAGT 1008
QY 834 CGCAAAGGAAGACAAGTTTATTTGGAGGTTATGACAAAGAAAGAAAGCAGTACGGCT 893
DB 1009 AGAAAAGGAAGACAAGTTTATCTGGAGGTTATGATATGGAGGAGAAAGCTGTCAGCA 1068
QY 894 TATGATTTAGCGGACTCAAGTATTTGGGAAACCACTACTACTACTACTACTACTACTACT 953
DB 1069 TATGATCTTCTGCACTCAAGTACTGGGTCCCTCTACTCACCACCAATTTCTCTGGGAG 1128
QY 954 GAATATGAAAAGAGGTAGAGAGATGAAGACATGACAGGCAAGAGTATGTCCTCA 1013
DB 1129 AATTATCAGAAAGAGATGAAGACATGAAGAAATGACTAGACAGAAATATGTCACAT 1188
QY 1014 CTGGCAGGAAAAGTACTGTTTCTCTCGGGTGCATGCAATTTATCGTGAGTAAACAAG 1073
DB 1189 TTGAGAGGAGGAGCAGTGGTTTCTCTAGGGGTGCTTCCATCTATAGAGGAGTCAACA 1248
|||||

```


QY	1074	CATCACCAATGGAAGATGGGAGCTAGGATAGGAGAGTCCCGGTACCAAGACCTC	1133
Db	1249	CATCACCAATGGAAGTGGGAGCTAGGATAGGAGAGTCCCGGTACCAAGATCTC	1308
QY	1134	TACTTGGGAATTTTGGCACACAAGAGAGCTGCAGAGGCATACGACATTCGGGCCATC	1193
Db	1309	TACTTGGGAATTTTGGAAACCAAGAGAGCTGCAGAGCTTACGATGTAGCAGCAATT	1368
QY	1194	AAATTCAGAGGATTAACCGCAGTACTTAATTCGACATGACATACAGATACCAAGCTTAAGCA	1253
Db	1369	AAATTCGCTGGCACAATGCTGTGACTTAACCTTTGATATCAGAGGTACGATGTTGTCGT	1428
QY	1254	ATCCT 1258	
Db	1429	ATCAT 1433	
RESULT 15			
LOCUS	AY080706		
DEFINITION	Arabidopsis thaliana putative ovule development protein		
ACCESSION	AY080706		
VERSION	AY080706.1	GI:19310586	
KEYWORDS	FUT CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 2056)		
AUTHORS	Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.		
TITLE	Arabidopsis Full Length cDNA Clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2056)		
AUTHORS	Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA		
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL CDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.		
The Salk, Stanford, PGE (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.			
Yamada, K. (SSP/PGE) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGE) contributed equally to this work as PIs.			
Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.			

FEATURES			
source	Location/Qualifiers		
	1..2056		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="4"		
	/clone="RAPL06-86-F1.9 (R11873)"		
	/ecotype="Columbia"		
	/note="This clone is in a modified pBluescript vector (lambda ZAP) as a XhoI/SstI insert."		
gene	1..2056		
	/gene="At4g37750"		
5'UTR	1..157		
	/gene="At4g37750"		
misc_difference	2		
	/gene="At4g37750"		
	/note="not present in genomic sequence"		
CDS	158..1825		
	/gene="At4g37750"		
	/codon_start=1		
	/evidence=experimental		
	/product="putative ovule development protein aintegumenta"		
	/protein_id="AAL85024.1"		
	/db_xref="GI:19310587"		
	/translation="MKSFCDNDNNHNTNLLGFLSSNNMKMGGRGGRRAIYSSSTSSAATSSVPPQVVGDNSTNFGVCGSNPGNGIYHSMVPLRSDGLCLMEALNRSSHNHDDSPKVEDEFGTHNNTSKHAMDLSLDLYNTTPEPTNTNFQFFSFPTNRHEETERNYCNGLSLTHGGSFNVVYGEFOQLSLMSPGSQSCQIVHSHQONONHOSHQOISEALVTSVGFPTTAAAKKKGQEDVWVVGOKIVHRSIDTFCQTSQYRGVTRHRTWGRVEAHLWDSNPKGHSRKGROVYLGQYMEKKAARAYDLAAKYGPSHTINFSAEYIOKEIEDNMNNTQYVAHLRRKSGFSKASISYRGVTRHAAKQWQARIRVAGNKDLYLGTQTEBAEAAYDVAALKFRGNVNTNFDITSDYDRIMSNTLLSGELARRNNSTIVRNTDQALNAVVEGSGNKEVSTTSPRLISFPFPAIFALPQVQKMFSGMGNMSPWTSNPNALKTVALTLPQMPVFAAWADS"		
	1826..2056		
	/gene="At4g37750"		
3'UTR			
ORIGIN			
Query Match	16.4%	Score 330.6	DB 8; Length 2056;
Best Local Similarity	75.4%	Pred. No. 4.9e-66;	
Matches 411;	Conservative 0;	Mismatches 134;	Indels 0; Gaps 0;
QY	714	AAACTATTGAGAGTTTGGACAGAGCGTCTATATACCGCGTGTACAGGCATCGG	773
Db	971	AAATCTATCGATATCTTTGGACACAGCACTTCTCAATACCGAGCGGTACAGACATAGA	1030
QY	774	TGGACAGCAAGATATGAGGCACATTTATGGGATATATAGTTGTAAGAGAGGCAACG	833
Db	1031	TGGACTGTAGATATGAAGTCATCTATGGGACATAGTTTCAAGAGGAAGGTACAGT	1090
QY	834	CGCAAGCAACAAAGTTTATTTGGAGGTTATGACAAAGAGAAAGAGCAGCTAGGCT	893
Db	1091	AGAAAAGAGACAAAGTTTATCTGGAGGTTATGATATGGAGGAGAAAGCTGCTGAGCA	1150
QY	894	TATGATTTAGCGGACATCAAGTATTGGGAAACACCACTACTACTACTCCCATGAGC	953
Db	1151	TATGATCTGTGCACCTCAAGTACTGGGTCCTCTACTCACCAATTTCTCTCGGAG	1210
QY	954	GAATATGAAAAGAGGTAGAGAGATGAGACATGACAAAGCAAGAGTATGTTGCCTCA	1013
Db	1211	AATTATCAAGAAAGATTGGAAGACATGAAGAACATGACTAGACAGAAATATGTTGCACAT	1270
QY	1014	CTGGCAGGAAAAGTAGTGGTTTCTCTGCTGCTCATCGATTATCGTGGAGTAAACAAGA	1073
Db	1271	TTGAGAGAGAGACAGTGGTTTCTTAGGGGCTCTCTATATAGAGAGTCAACAAGA	1330
QY	1074	CATCACCACATGGAAGATGCAAGCTAGGATAGAGAGTCGCCGCTAACAAAGACCTC	1133
Db	1331	CATCACCAGCTGAAGGTGCAAGCAGGATTTGGTAGCTGCTGGAACAAAGATCTC	1390
QY	1134	TACTTGGAACTTTTGGCAGACAGAGAGCTCAGAGGATACGATTCGCGCCATC	1193
Db	1391	TACCTTGGAACTTTTGGAAACCAAGAGAGCTGAGAGGCTTACGATGTAGCACT	1450

Oy	1194	AAATTCAGAGGATTAAACCGCAGTGA	CTAACTTCGACATGAACAGATACAACGTTAAAGCA	1253
Db	1451	AAGTTCCTGGCACAATGCTGTGACTAAC	TTTGATATCAGAGGTACGATGTTGATCGT	1510
Oy	1254	ATCCT	1258	
Db	1511	ATCAT	1515	

Search completed: July 19, 2004, 08:51:46
Job time : 7912 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2004, 05:17:49 ; Search time 5287 Seconds
(without alignments)
11375.531 Million cell updates/sec

Title: US-09-980-364-1

Perfect score: 2014
Sequence: 1 gttctctctctctctttaag.....gttgataaaatttcagact 2014

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_esth:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	534.4	26.5	556	B2512981	B2512981 BOMQP05TF
c 2	393.4	19.5	781	BH443187	BH443187 BOGVV76TF
c 3	382.6	19.0	769	CA783156	CA783156 sac20d05
4	370.2	18.4	720	BQ864461	BQ864461 QGC26M12

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	361.4	17.9	590	12	BM094116
6	351.4	17.4	748	14	CF807326
7	346.4	17.2	674	10	BF647766
8	344.4	17.1	552	13	BU964897
9	339.4	16.9	644	10	AW200688
10	339.2	16.8	882	14	CK267021
11	335.4	16.7	500	10	AW780688
12	335.2	16.6	593	12	BM307997
13	333.6	16.6	717	28	BZ044547
14	332	16.5	686	28	BZ496117
15	323.6	16.1	545	14	CD475882
16	317.2	15.7	585	12	BM086088
17	308.6	15.3	631	13	CA103041
18	307.4	15.3	640	13	CA100375
19	307.2	15.3	697	14	CA189006
c 20	306	15.2	481	28	BH694176
21	305	15.1	555	12	BJ188928
22	305	15.1	558	12	BJ178045
23	304.2	15.1	690	13	CA094556
24	302	15.0	641	14	CD879292
25	300.4	14.9	518	28	BH443196
26	299.8	14.9	1753	11	AY109146
27	297.2	14.8	663	28	BH927892
28	296.6	14.7	767	28	BZ077316
29	288.4	14.3	466	13	BU965263
c 30	273	13.6	417	28	BH908835
31	272.6	13.5	599	14	CA230632
32	271.2	13.5	1160	14	CK206573
33	270.8	13.4	558	14	CA232734
34	268.2	13.3	476	12	BM092890
35	268.2	13.3	545	12	BI974354
36	265	13.2	527	12	BM309412
37	256.4	12.7	432	12	BG882497
38	254.4	12.6	596	12	BM307877
39	253.4	12.6	642	28	BZ516982
40	251.4	12.5	513	13	BQ122372
c 41	251.2	12.5	497	14	CF244784
42	248.4	12.3	842	14	CF446573
43	245.6	12.2	704	13	BQ625052
44	244.6	12.1	469	14	CF324329
45	242.4	12.0	420	9	AJ475492

ALIGNMENTS

RESULT 1
BZ512981/c BZ512981 556 bp DNA linear GSS 16-DEC-2002
LOCUS BOMQP05TF BO_2_3_KB Brassica oleracea genomic clone BOMQP05,
DEFINITION genomic survey sequence.
ACCESSION BZ512981
VERSION BZ512981.1 GI:27039560
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea.

REFERENCE
AUTHORS Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other_GSSs: BOMQP05TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: Tf
Class: Sheared ends.

Brassicaceae; Brassica.
rosids; eurosids II; Brassicales; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
euosids II; Brassicales; Streptophyta; Embryophyta; Tracheophyta;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 769)

Shoemaker, R., Keim, P., Vodkin, L., Expelding, J., Corvett, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, I., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, J., Gibbons, M., Pape, D., Harvey, N., Schurr, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 454.

FEATURES

Location/Qualifiers

1..769

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl036-14361"

/tissue_type="somatic embryos cultured on MSD 20"

/lab_host="DH10B"

/clone_lib="Gm-cl036"

/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Query Match 19.0%; Score 382.6; DB 14; Length 769;

Best Local Similarity 76.6%; Pred. No. 2.1e-66;

Matches 482; Conservative 0; Mismatches 144; Indels 3; Gaps 1;

QY 692 TAGCGTTGAAGCTACACCGAGAAACTATTGAGCTTTGGACAGACGCTCATATA 751

DB 94 TACTGTTGAATAGCACTTAGAAGAACTTTGGATACATTCGGCGAGAACATCCATATA 153

QY 752 CCGCGGTGTTTACAAGGCATCCGTGGACGAGGAATATGAGGCACATTTATGGGATAATAG 811

DB 154 TCGTGGAGTAACTCGACATAGATGAGCTGGAAGGTATGAGCTCATCTTTGGGATAATAG 213

QY 812 TTGTAAAGAGAGGCGCAACCGCGCAAGGAAGACAGAGTTTATTTGGGAGTTATGACAA 871

DB 214 CTGTAGAAGGGAAGGCGCAATCAAGAAAGGACGCCAAGTTTATTTGGGTGGATATGATA 273

QY 872 AGAAGAAAAGCAGCTAGGCTTATGATTTAGCGGCACTCAAGTATTTGGGAAACCAACAC 931

DB 274 AGAAGAGAAGCAGCTAGACCTTATGATTTAGTCTCTCTGAACTACTGGGGACATCCAC 333

QY 932 TACTACTACTTCCCATGAGCGGAATATGAAAAGAGGTAGAGAGATGAGACATGAC 991

DB 334 CACTTACCACATTTCCCAATTAGCAACTATGAGAAGGAATTTGGATGCAATGAAACACATGAC 393

QY 992 AGGCGAAGAGTATGTTGCTCTCACTGCGCAGGAAGAAAGTAGTGGTTCTCTCGTGGTGCATC 1051

DB 394 GAGACAAGAATTTGTTGCGCGCAATTAGAAAGGAAGAAAGCAGTGGTTCTCCAGGGGTGCATC 453

QY 1052 GATTATTCGTGGAGTAAACAAGACATCACCAACATGGAAGATGCGCAAGCTAGGATAGGAAG 1111

DB 454 AATGTATCGTGGAGTTTACAAGGCATCACCAACACGGAAGATGCGCAAGCAAGGATTGGCAG 513

QY 1112 AGTCGCGGTAAACAAGACCTCTTCTTGGAACTTTTGGACATTTTGGCACACAGAAGAGCTGCAGA 1171

DB 514 AGTTGCGAGGAACAAAGATCTTTTACTTGGGAATTTTCACTAGTACTGAGGAAGAGCTGCAGA 573

QY 1172 GGCATACGACATTTGCGGCCATCAAAITTCAGAGGATTAACCGCAGTCACTAACTTCGACAT 1231

DB 574 AGCATACGACATAGCAGCGGATAAAGTTCAGAGGTCTCAACGCTGTCAACAATCTTCACAT 633

QY 1232 GAACAGATACACGTTTAAAGCAATCTCGAAGCCCTAGTCTTCCTAT---TGTGAGCGC 1288

DB 634 GAGCCGATACGAGTGAAGCCATTTCTGAAAGCAACACTCTCCCAATAGGAGGAGCGCGC 693

QY 1289 CGAAGAACGCTCAAGGAGGCTAACCGTC 1317

DB 694 TGCAGCGCTCTGAAGAGCTCAAGCTC 722

RESULT 4

BO864461 720 bp mRNA linear EST 14-AUG-2002

LOCUS QGC26M12.yg.ab1 QG_ABCDI lettuce salinas Lactuca sativa cDNA clone

DEFINITION QGC26M12, mRNA sequence.

ACCESSION BO864461

VERSION BO864461.1 GI:22249926

KEYWORDS EST.

SOURCE Lactuca sativa

ORGANISM Lactuca sativa

REFERENCE Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae; Lactuca.

AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.

TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project

JOURNAL http://compgenomics.ucdavis.edu/

COMMENT Unpublished (2002)

Contact: Alexander Kozik [R.W.Michelmore]

Department of Vegetable Crops, R.W.Michelmore Lab

University of California at Davis (UCD)

Asmundson Hall, UCD, Davis, CA 95616, USA

Tel: 1-(530)-742-1742

Fax: 1-(530)-752-9659

Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]

singleton, see http://cgdb.ucdavis.edu/ for details.

Plate: QGC26 row: M column: 12.

FEATURES

Location/Qualifiers

1..720

/organism="Lactuca sativa"

/mol_type="mRNA"

/cultivar="Salinas"

/db_xref="taxon:4236"

/clone="QGC26M12"

/lab_host="E.coli"

/clone_lib="QG_ABCDI lettuce salinas"

/note="Vector: pBRCNASTAB; the library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize

size bias. Details of each source of RNA and library construction can be obtained at <http://cspdb.ucdavis.edu/>.
TAG_TISSUE=chemical induction
TAG_LIB=QG ABCDI lettuce salinas
TAG_SEQ=TGTAGCGGG

ORIGIN

```

Query Match      18.4%; Score 370.2; DB 13; Length 720;
Best Local Similarity 79.5%; Pred. No. 6.6e-64;
Matches 438; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

783 AGATATGAGGACATTTATGGGTAATAGTTGTAAGAAGAGGCGCAAGCGCAAGGA 842
    |||
Db      2 AGATACGAGGCTCATTTATGGGCAATAGTTGCGAAGAAGAGCAAGTTCGAAGGA 61
    |||

783 AGACAAGTTATTTGGGAGTTATGACAAAGAAGAAAAGCAGTACGGCTTATGATTTA 902
    |||
Db      62 AGACAAGTTATTTGGGTTAGCAAAAGAAAGATAAGGAGCAAGAGCTTATGATTTA 121
    |||

903 GCCGCACTCAAGTATTCGGGAACCACTACTACTACTTCCCATGACGCGAATATGAA 962
    |||
Db      122 CTTGCATTTGAATTTGGGTACCAACTACGCAAAATTTCCCTATTACCAACTATGAG 181
    |||

963 AAAGAGTAGAAGAGATGAAGCAGCATGACAAGGCAAGAGATGTTGCTCTCACTGCGCAGG 1022
    |||
Db      182 AAGGAGATCGAGGAGATGAACACATGACTAGACAAAGAAATACGTAGCATCACATTAGAGG 241
    |||

1023 AARAGTAGTGGTTCTCTCGTGGGCGATCATTTATGCTGAGTACAAAGACATCACCAA 1082
    |||
Db      242 AAAAGTAGCGGTTTTCTCGAGGTGCTTCATCTATAGAGGTGTGACAGACACCATCAA 301
    |||

1083 CATGGAAGATGGCAAGCTAGGATAGGAAGAGTGGCGGTAACAAAGACCTCTACTTTGGGA 1142
    |||
Db      302 CATGGAAGATGGCAAGCAAGAAATTTGAAGAGTTGCGAGGAAACAAAGATCTTTACTTTGGGA 361
    |||

1143 ACTTTTGGCACAAAGAAGAGCTGAGAGGCATACAGATTCGGGCGCATCAAAATTCAGA 1202
    |||
Db      362 ACTTTTGGTACACAGGAAGAGCTGCTGAGGCTTATGACATTCGTCGCATCAAAATTCGGT 421
    |||

1203 GGATTAACCGCAGTGAATCTCGACATCAAGACAGATACAAAGCTTAAAGCAATCTCTCGAA 1262
    |||
Db      422 GGACTGATGAGTCAACAACTTTGAATCAACAGATACGACGTTAAATGATCACTTGA 481
    |||

1263 AGCCCTAGTCTTCTATGTTAGCGGCGCAAAAGCTCTCAAGGAGGCTAACCGTCGCGTT 1322
    |||
Db      482 AGCACCACCTTACCGGTTGGTGTGCGAGTAAACGCTCAAGGATGCAAGAACGCGCT 541
    |||

1323 CCAAGTATGAT 1333
    |||
Db      542 GCNACTGAGTT 552
    |||

```

RESULT 5
BM094116
LOCUS
DEFINITION
sah25f05.y1 Gm-cl036 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl036-2554 5' similar to TR:Q9SYC2 Q9SYC2 F11M15.6 PROTEIN.
// mRNA sequence.
BM094116
VERSION
KEYWORDS
SOURCE
ORGANISM
Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustroids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 590)

REFERENCE
AUTHORS
Shoenaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,

TITLE
JOURNAL
COMMENT

McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 431.

FEATURES
source

Location/Qualifiers
1..590
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl036-2554"
/tissue_type="somatic embryos cultured on MSD 20"
/lab_host="DH10B"
/clone_lib="Gm-cl036"
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This
cDNA library was constructed from mRNA isolated from
somatic embryos (age ranging from 2 months to 9 months)
cultured on MSD 20. The library was prepared using the
Life Technologies Superscript cDNA library construction
kit. Complementary DNA was synthesized from mRNA using a
poly (dT) sequence with a NotI restriction site. SalI
linkers adapters were ligated to the blunt-ended cDNA
fragments followed by NotI digestion. The cDNA fragments
were directionally cloned into the NotI-SalI restriction
site of the pSPORT1 vector. The ligated cDNA fragments
were transformed into E.coli ElectroMax DH10B host cells.
This library was constructed in the laboratory of Dr. Lila
Vodkin by Anu Khanna at the University of Illinois at
Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

```

Query Match      17.9%; Score 361.4; DB 12; Length 590;
Best Local Similarity 77.2%; Pred. No. 4e-62;
Matches 453; Conservative 0; Mismatches 131; Indels 3; Gaps 1;

729 TTTCGACAGAGGAGCTCTATATACCGCGTCTTACAAGGCATCGTGGCAGGAAGATAT 788
    |||
Db      4 TTTCGACAGAGGAGCAATCCATATATCTGTTGAGTAAACGACATAGATGATGAGGTAT 63
    |||

789 GAGGCACATTTATGGGATATAGTTGTAAAGAGAGCGCCAAAGCGCAAGAGAACAA 848
    |||
Db      64 GAAGCTCACCTTTGGGATATAGCTGTAGAGGAGGCAATCAGAAAAGGACGCCAA 123
    |||

849 GTTTATTGGGAGGTATGACAAAGAGAAAAGCAGCTAGGCTTATGATTTAGCCGCA 908
    |||
Db      124 GTTTATTGGGAGGTATGATTAAGAAAGAAAAGCAGCTAGGCTTATGATTTAGCTGCA 183
    |||

909 CTCAGATTTGGGAAACCCACCTACTACTTCCCATGAGCGAATATGAAAAGAG 968
    |||
Db      184 CTGAAGTACTGGGGACATCCACCCTTCCCAATTTAGTAATCTATGAGAAGAA 243
    |||

969 GTAGAAGAGATGAGCAGACATGACAGGCAAGAGATGTTGCTCTCACTGCGCAGGAAAGT 1028
    |||
Db      244 TTGATGAATGAACACATGACGCGCAAGAAATTTGTTCTGCCATTAGAAGAAAGC 303
    |||

1029 AGTGGTTCTCTCGTGTGATCGAATTTATCGTGGAGTAAACAGACATCACCAACATGGA 1088
    |||
Db      304 AGTGGTTCTCTCGAGGTGATCAATGATATCGTGGAGTTACAAGGCATCACCAACACGGA 363
    |||

1089 AGATGCGCAAGCTAGGATAGAGAGTCCGCGTGAACAGACCTCTACTTGGAACTTTT 1148
    |||
Db      364 AGATGCGCAAGCAAGATTTGCGAGGTGCGAGGAACAAAGATCTTTACTTTGGAACTTTC 423
    |||

1149 GGCACACAGAGAGCTGCGAGAGGCATACGACATTTGCGGCCATCAAAATTCAGAGGATTA 1208
    |||

```

```

Db      424 AGTACTGAAGAGAGGCTGCTGAAGCATACACATAGCTGCGATAAAGTTGAGAGTCTC 483
QY      1209 ACCGCACTGACTTAATCTCGACATGAACAGATACACAGTTAAGCAATCTCGAAAGCCCT 1268
Db      484 AACGCTGTACAAACTTTGACATGAGCGCTACGAGGTGAAGCCATCTTTGAAAGCAAC 543
QY      1269 AGTCTTCTCTAT---TGGTAGCGCCGCAAAAGCTCTCAAGGAGGCTAA 1312
Db      544 ACTCTCCCAATAGGAGGAGGAGCTGCAAAAGCGTCTGAAGAAGCTCA 590

RESULT 6
CF807326
LOCUS   CF807326
DEFINITION psHB023xP07f USDA-IRAFs:Expression of Phytophthora sojae genes during infection and propagation Phytophthora sojae cDNA clone sHB023P07 5, mRNA sequence.
ACCESSION CF807326
VERSION    CF807326.1 GI:37995737
KEYWORDS  EST.
SOURCE    Phytophthora sojae
ORGANISM  Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.
REFERENCE 1 (bases 1 to 748)
AUTHORS  Tyler,B.
TITLE    Tyler,B. Not Published
JOURNAL  Unpublished (2003)
COMMENT  Contact: Tyler B
          Tyler lab
          VBI
          1880 Pratt Dr., Blacksburg, VA 24061, USA
          Tel: 540-231-7318
          Email: bmtylev@vt.edu
PCR PRIMERS
FORWARD: BK reverse primer
BACKWARD: BK reverse primer
Plate: 023 row: F column: 07
Seq primer: BK reverse primer
High quality sequence stop: 748.
          Location/Qualifiers
FEATURES             source
1..748
/organism="Phytophthora sojae"
/mol_type="mRNA"
/db_xref="taxon:67593"
/clone="sHB023P07"
/tissue_type="mycelium"
/cell_line="P6497"
/dev_stage="48 hr. post infection stage"
/lab_host="Soybean plant"
/clone_lib="USDA-IRAFs:Expression of Phytophthora sojae genes during infection and propagation"
/note="Vector: pBK-CMV, Site_1: EcoR1; Site_2: Xho1"

ORIGIN
Query Match 17.4%; Score 351.4; DB 14; Length 748;
Best Local Similarity 77.5%; Pred. No. 3.9e-60;
Matches 424; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY      700 AAGCTACACCGAAGAAATATTGAGAGTTTGGACAGAGGAGCTCTATATACCGCGGTG 759
Db      202 AACCTCTCCAAAGAAACCGTCGACACCTTCGCGCAACGCACCTCCATCTACCGCGGCG 261
QY      760 TTCAAGGCATCGGTGACAGAGAGATATAGGCACATTATTTGGGATATAGTTGTAAAA 819
Db      262 TCACCCGACATAGATGACGCGGAAGATACCAAGCTCATCTATGGGACAATAGTTGTAGAA 321
QY      820 GAGNAGCCCAACCGCAAGAGGAGCAAGCTTTATTTGGGAGGTATGACAAGAGAGAAA 879
Db      322 GAGAAGGCCCAAGCAGGAAGGAGCAAGTTTACCTGGTGTGTTTNGACAAGAGACATA 381
QY      880 AAGCAGCTAGGCGTTATGATTATAGCCGCACTCAAGTATTGGGGAACCAACCACTACTACTA 939

```

```

Db      382 AGCAGCCAGGGCTTACGATCTCGCAGCTCTCAAGTACTGGGTCCAACTACCACCACCA 441
QY      940 ACTTCCCATGAGCGGAATATGAAAGAGGTAGAGATGAGACATGACAGCAAGCAAG 999
Db      442 ACTTCCCATTTCCAACTATGAGAGGAACTGGAGAGATGAAGAAACATGACAGGCAAG 501
QY      1000 AGTATGTGCTCCTCAGTCGCGCAGGAAAAGTAGTGTCTCTCTCGTGGTGCATCGAATTATC 1059
Db      502 AGTTGTGTGCTTCTCTACGAAGNAGAGCAGTGTCTCTCTAGGGGGGCTCTATATACA 561
QY      1060 GTGAGGTAAACAAGACATCACCAACATGGAAGATGGCAAGCTAGGATAGGAAGAGTGCCTG 1119
Db      562 GAGGAGTGCAGACACACCACGATGGCGGATGGCAGGCGAGATAGGCAGAGTTGCGG 621
QY      1120 GTAAACAAGACCTCTACTTTGGGAACCTTTGGCACACAAGAAAGCTGCAGAGGCATACG 1179
Db      622 GAAACAAGAACCTCTACTCTTGGAACTTTTACGACCCCAAGAGAGCTGCTGAGGCCCTATG 681
QY      1180 ACATTGGCGGCATCAAAATTCAGAGGATTAACCCAGTACTAACTTCGACATGAACAGAT 1239
Db      682 ACATTGCTGTATCANATTCAGGGGATTAATGCACTCANCAACTTTGACATGATCGCT 741
QY      1240 ACAAGCT 1246
Db      742 ACGATGT 748

RESULT 7
BF647766
LOCUS   BF647766
DEFINITION BF647766 674 bp mRNA linear EST 20-DEC-2000 clone NF025G09EC 5', mRNA sequence.
ACCESSION BF647766
VERSION    BF647766.1 GI:11912896
KEYWORDS  EST.
SOURCE    Medicago truncatula (barrel medic)
ORGANISM  Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
REFERENCE 1 (bases 1 to 674)
AUTHORS  Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
TITLE    Expressed Sequence tags from the Samuel Roberts Noble Foundation - Center for Medicago Genomics Research
JOURNAL  Unpublished (2000)
COMMENT  Contact: Dixon RA
          Plant Biology Division
          The Samuel Roberts Noble Foundation
          2510 Sam Noble Parkway, Ardmore, OK 73402, USA
          Tel: 580 221 7302
          Fax: 580 221 7380
          Email: radixon@noble.org
          Insert Length: 674 Std Error: 0.00
          Plate: 025 row: G column: 09
          Seq primer: TCACACAGGAACACGCTATGAC.
FEATURES             source
1..674
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF025G09EC"
/tissue_type="Cell cultures derived from root tissues"
/dev_stage="Cell suspensions were subcultured every 14 days. Cells were induced six days after subculture"
/clone_lib="Elicited cell culture"
/note="Vector: Lambda Zap; Cells were induced with yeast cell wall extracts equivalent to 50ug/ml glucose in the final concentration. Samples were taken at 0.5, 1, 12 and 24 hours after induction. Equal amounts of RNA from each time point were pooled and used for mRNA isolation."

```

ORIGIN

Query Match 17.2%; Score 346.4; DB 10; Length 674;
 Best Local Similarity 81.5%; Pred. No. 3.9e-59;
 Matches 401; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 815 TAAAGAGAGCCCAAGCCCAAGGAGCAAGTATTTGGAGGTTATGACAAAGA 874
 Db 1 TAGAGAGAGGGGAGAGCTCCCAAGGAGGCAAGTTTACTTGGAGGTTATGACAAAGA 60

QY 875 AGAAAAAGCAGTAGGGCTTATGATTTTAGCCGCACTCAAGTATTGGGGAACCACTAC 934
 Db 61 AGAAAAAGCAGTAGAGGCTATGATTTGGGAGCACTAAATATTTGGGGAACCACTACTAC 120

QY 935 TACTAATCTCCCATGAGCGCAATATGAAAAAGAGGTAGAAGATGAAGCACATGACAAG 994
 Db 121 AACAAATTTCCAAATTAGCCATTATGAAAAAGAGTGGAAAGATGAAGCATATGACAAG 180

QY 995 GCAAGAGTATGTCCTCAGTGCAGGAGAAAGTAGTGTCTCTCGTGGTGCATCGAT 1054
 Db 181 GCAAGAGTATGTCCTCAGTGCAGGAGAAAGTAGTGTGTCTCTCGTGGTGCATCGAT 240

QY 1055 TTATCGTGGATTAACAAGACATCAACATGGAAGATGGAAGTGAAGTAGGATAGGAAGAGT 1114
 Db 241 TTATCGGAGGAGTAAACAAGACATCAACATGGAAGTGAAGTGAAGTGAAGAGT 300

QY 1115 CGCCGGTAAACAAGACCTTACTTTGGAACTTTGGCACAACAAGAGAGCTCGAGGC 1174
 Db 301 TGCAGGCAACAAGATCTTACCTAGGAATCTTCAGCACTCAAGAAGAGGACGACAGGC 360

QY 1175 ATACGACATTCGGGCCATCAAAATTCAGAGGATTAACCGCAGTCACTAACTTCGACATGAA 1234
 Db 361 ATATGATGTGCAGCAATAAAATTCAGAGGATTCAGGATTCAGAACTTTGACATGAG 420

QY 1235 CAGATCAAGGTTAAAGCAATCTCGAAAGCCCTAGTCTTCTTATGTTGGTAGCCGCAAA 1294
 Db 421 CAGATATGATGTCAAAACCATCTTGAAGCAGCAGCATTAACCAATTTGGTGTCTGCAAA 480

QY 1295 ACGTCTCAAGGA 1306
 Db 481 GCGTTTAAAGA 492

RESULT 8

BU964897
 LOCUS sat04a07.v1 Gm-c1036 Glycine max cDNA clone SOYBEAN CLONE ID:
 DEFINITION Gm-c1036-12685 5' similar to TR:Q9SYC2 Q9SYC2 FILM15.6 PROTEIN. ;
 mRNA sequence.

ACCESSION

BU964897
 VERSION BU964897.1 GI:24205644

KEYWORDS

EST.
 SOURCE Glycine max (soybean)

ORGANISM

Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

REFERENCE

1 (bases 1 to 552)
 Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvett, V.,
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
 Beck, C., Wylie, T., Underwood, K., Streptoe, M., Theising, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
 Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Willson, R.
 Public Soybean EST Project
 Unpublished (1999)

TITLE

JOURNAL Public Soybean EST Project

COMMENT Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact: ccu@resgen.com web site:
 www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 445.

Location/Qualifiers

1. 552

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-c1036-12685"

/tissue type="somatic embryo" cultured on MSD 20"

/lab host="DH10B"

/clone lib="Gm-c1036"

/note="Vector: PSORT1; Site 1: NotI; Site 2: SalI; This

cDNA library was constructed from mRNA isolated from

somatic embryos (age ranging from 2 months to 9 months)

cultured on MSD 20. The library was prepared using the

Life Technologies SuperScript cDNA library construction

kit. Complementary DNA was synthesized from mRNA using a

poly (dT) sequence with a NotI restriction site. SalI

linkers adapters were ligated to the blunt-ended cDNA

fragments followed by NotI digestion. The cDNA fragments

were directionally cloned into the NotI-SalI restriction

site of the pSPORT1 vector. The ligated cDNA fragments

were transformed into E.coli ElectroMax DH10B host cells.

This library was constructed in the laboratory of Dr. Lila

Vodkin by Anu Khanna at the University of Illinois at

Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Query Match 17.1%; Score 344.4; DB 13; Length 552;
 Best Local Similarity 77.5%; Pred. No. 1e-58;
 Matches 417; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 754 GCGGTGTTCACAGGCACTCGGTGGACAGGAGATATGAGGACACATTTATGGGATATAGTT 813
 Db 1 GTGGAGTAAACCGACATAGATGGACTGGAGGTATGAGCTCACCTTTGGATATAGCT 60

QY 814 GTAAAGAGAGGCCAAACCGCCAAAGCAAGATTTATTTGGGAGGTTATGACAAAG 873
 Db 61 GTAGAAGGGAAGGCAATCAAGAAAGGAGCGCCAAAGTTTATTTGGGTGGATATGATAAG 120

QY 874 AAGAAAGCAGCTAGGCTTATGATTTAGCCCACTCAAGTATTGGGGACCACTA 933
 Db 121 AAGAAAGCAGCTAGGCTTATGATTTAGCTGCACTGAAGTACTTGGGGACATCCACCA 180

QY 934 CTACTAACTTCCCATCAGCGCAATATGAAAAGAGGTTAGGAAGATGAAGCACATGACAA 993
 Db 181 CTACCAACTTTCCCAATTAGTAACATATCAGAAGGAATTTGGATGAAATGAACACATGACGC 240

QY 994 GCGAAGTATGTTGCTCTACTCGCAGGAAAGTAGTGTCTTCTCTCGTGGTGCATCGA 1053
 Db 241 GACAAAGATTTGTTGCTGCCATTAGAAGGAAAGCAGTGTGTTCTCCAGGGGTGCATCAA 300

QY 1054 TTTATCGTGGAGTAAACAAGACATCAACCAATGGAAGTGGCAAGCTAGGATAGGAAGAG 1113
 Db 301 TGTATCGTGGAGTAAACAAGACATCAACCAACAGGAGATGCGCAAGATTTGGCAGAG 360

QY 1114 TCGCCGGTAAACAAGACTCTACTTGGGAATTTTGGCACAAGAGAAAGCTGCGAGAG 1173
 Db 361 TTGAGGAAACAAGATCTTTACTTGGGAATTTTCACTACTGAAGAGAGGCTGCTGAAG 420

QY 1174 CATACGACATTTGGGCGCATCAAAATTCAGAGGATTAACCGCAGTGCATTAACCTCGCATGA 1233
 Db 421 CATACGACATGCTGCGATTAAGTTTCAAGGCTCTCAACGCTGTCAAAACTTTGACATGA 480

QY 1234 ACAGATCAACAGTTAAAGCAATCTCGAAAGCCCTAGTCTTCTATTTGGTAGCCGCGC 1291
 Db 481 GCGCTACGACGTGAAGAGCCATCTCTTGAAGCAACACTCTCCCAATAGGAGGAGGAGC 538

RESULT 9
 LOCUS
 DEFINITION

AW200688 644 bp mRNA linear EST 02-DEC-2001
 se92c07.y1 Gm-cl027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 Gm-cl027-229 5' similar to TR:Q41832 Q41832 OPEN READING FRAME. ;
 mRNA sequence.

AW200688
 AW200688.1 GI:6481417
 EST
 KEYWORDS
 SOURCE
 ORGANISM

Glycine max (soybean)
 Glycine max
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 644)
 Shoemaker,R., Keim,P., Vodkin,L., Expelding,J., Coryell,J.,
 Khanna,A., Bolla,B., Marra,M., Hillier,D., Kucaba,T., Martin,J.,
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
 McCann,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estwatson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccu@resgen.com
 Insert Length: 968 Std Error: 0.00
 High quality sequence stop: 418.
 Location/Qualifiers
 1..644
 /organism="glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl027-229"
 /tissue_type="cotyledons of 3- and 7-day-old Williams
 seedlings"
 /lab_host="DH10B"
 /clone_lib="Gm-cl027"
 /notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
 XhoI; This cDNA library was constructed from mRNA isolated
 from cotyledons of 3- and 7-day-old Williams seedlings
 which were propagated on paper towels with distilled
 water. The cotyledons were flash-frozen in liquid
 nitrogen, then lyophilized for 72 hours. Unequal amounts
 of mRNA was used for cDNA synthesis. Stratagene's cDNA
 Synthesis Kit (catalog number 200401) was used to
 synthesize the cDNA. First- strand synthesis was
 performed with 5-methyl dCTP, hence the ligated cDNA was
 hemimethylated. A modification of Stratagene's
 first-strand synthesis primer was used. An anchor
 nucleotide (V=A, C, or G) was added to the 3' end of the
 primer [GAGACAGAGAGAGAGAGACTAGCTCGAG(T)18] to anchor
 the primer at the 5' end of the poly(A) tract. After
 second- strand synthesis, the cDNA ends were filled in
 with cloned Pfu DNA, ligated to EcoRI adapters and
 subsequently phosphorylated. The XhoI site within the
 first-strand synthesis primer was then restricted by
 digestion with XhoI; all XhoI sites in the cDNA would be
 protected by their hemimethylated status. The cDNA
 constructs were size-fractionated with a 500 bp cutoff,
 using GibcoBRL Life Technologies' cDNA Size Fractionation
 column. The column eluent was then ligated into
 Stratagene's pBluescript(tm) II XR predigested vector
 (pBluescript II SK+), that has been digested with EcoRI
 and XhoI.

FEATURES
 source

ORIGIN

Query Match 16.9%; Score 339.4; DB 10; Length 644;
 Best Local Similarity 77.8%; Pred. No. 1e-57;
 Matches 409; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 700 AAGCTACACGAGAGAAACTATTGAGAGTTTGGACAGAGGACGCTATATACCGCGGTG 759
 Db 118 AACCTTCTCAAGAGAAACCGTCGACACCTTGGCCACGACCTCCATCTACCGGCG 177
 QY 760 TTCAAGGCGCATCGGTGGACAGGAGATATAGCGCACTTATPGGATATATAGTTGTA 819
 Db 178 TCACCCGACATAGATGACGCGGAAGATACGAAGCTCATCTATGGGACAAATAGTTGTAGAA 237
 QY 820 GAGAGGCGCAACGCGCAAGAGAGACAAGTTTATTTGGGAGGTTATGACAAAGAGAAA 879
 Db 238 GAGAAGGCCAAAGCAGGAAAGGAGCAAGTTTACCTGGTGGTTATGACAGAGAGATA 297
 QY 880 AAGCAGCTAGGGGCTTATGATTATTCGCGCACTCAAGATTATTTGGGGAACCACTACTACTA 939
 Db 298 AGGAGCGCCGGCTTACGATCTCGAGCTCTCAAGTACTGGGTCCAACTACCAACCA 357
 QY 940 ACTTCCCATGAGGGAATATGAAAGAGGTAGAGAGATGAAGCACATGACAAGGCAAG 999
 Db 358 ACTTCCCATTTCCAACTATGAGAAGGAACTGGAGAGATGAAGAACATGACAGGCAAG 417
 QY 1000 AGTATGTTGCTCACTGCGCAGGAAAAGTAGTGGTTTCTCTGCTGTCATCGATTATTC 1059
 Db 418 AGTTTGTGCTTCTTACGAGAGAGAGCAGTGGTTCTTAGGGGGGCTCTATATACA 477
 QY 1060 GTGAGTAAACAAGACATCACCAATGGAAGATGGCAAGCTAGGATAGGAAGAGTCGCG 1119
 Db 478 GAGAGTGAGGAGACACCAACGAGCATGGCGGATGGCAGCGAGAATAGCAGAGTTGCGG 537
 QY 1120 GTACAAGACCTCTACTTGGGAACCTTTTGGCACACAAGAGAGCTGCAGAGGCATAGC 1179
 Db 538 GAACAAGACCTCTACTTGGGAACCTTTTGGCACACAAGAGAGCTGCAGAGGCATAGC 597
 QY 1180 ACATTGGCGGCATCAAAATTCAGAGGATTAACCGCAGTACTAATT 1225
 Db 598 ACATTGCTGTATCAAAATTCAGGCGATTAATTCAGTGCACAAACCT 643

RESULT 10
 LOCUS
 DEFINITION

CK267021 882 bp mRNA linear EST 12-DEC-2003
 ES7713099 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POAC822 5' end, mRNA sequence.

CK267021
 CK267021.1 GI:39823999
 EST
 KEYWORDS
 SOURCE
 ORGANISM

Solanum tuberosum (potato)
 Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 882)
 Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
 Generation of ESTs from abiotic stressed potato tissue
 Unpublished (2003)
 Other ESTs: EST713100
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.
 Location/Qualifiers

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES

```

source
1. .882
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAC822"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="potato abiotic stress cDNA library"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d,
and 2d). Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

ORIGIN
Query Match 16.8%; Score 339.2; DB 14; Length 882;
Best Local Similarity 72.2%; Pred. No. 1e-57;
Matches 456; Conservative 0; Mismatches 173; Indels 3; Gaps 1;

QY 673 AAGGAGAGACTATTGATGATAGCGTTGAAGCTACACCGAAGAAACTATTGACAGTTTG 732
Db |||||
QY 116 AAAAGGCGATTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175
Db |||||
QY 733 GACAGAGACGCTCTATATACCGCGTGTTCACAGGCTCGGTGCACAGGAAGATATGAGG 792
Db |||||
QY 176 GGCAAGAACTTCAATCTACAGAGGTGTACAGACATAGTACAGAGGAAGATATGAAG 235
Db |||||
QY 793 CACATTTATGGGATAATAGTTGTAAAGAGAGCGCCAAAGCGCAAGGAAGCAAGTTT 852
Db |||||
QY 236 CTCATCTATGGGATAACAGCTGTAGGAGAGAGGCGCAAGCTAGAAAAGGCGCTCAAGTGT 295
Db |||||
QY 853 ATTTGGCAGGTTATGACAAAGAGAAAAGCAGCTAGGCGTGTATGTTTAGCGGCACTCA 912
Db |||||
QY 296 ACTTGGTGGATATGACAAAGAGATATAGCAGCAAGGCGATATGACTTGGCAGCTCTTA 355
Db |||||
QY 913 AGTATTGGGAAACCACTACTACTTAATCTCCCATAGCGAATATGAAAAAGAGGTAG 972
Db |||||
QY 356 AGTATTGGGTTCCACAGCTACCACTCCCTGTTACTAATATTAATAAGAAATTCG 415
Db |||||
QY 973 AAGAGATGAAGCATCAGACAGGCAAGATATGTTCCTCCTCAGTCCGCGAAGAGTAGTG 1032
Db |||||
QY 416 AGGAATGAACACATGACTATAGCAAGAAATCTATGCTCTTAAGAGAAAGTAGTG 475
Db |||||
QY 1033 GTTCTCTCGTGTGTCATCGATTATCGTGGAGTAAACAGACATCAACATCGGAAGAT 1092
Db |||||
QY 476 GTTCTCTCCCGAGGCTTCGATTACCGGGTGTGACAGGCATCATCAACAGGCGGCT 535
Db |||||
QY 1093 GGAAGCTAGGATAGGAGATCGCGGTAAACAGACTCTACTTTGGGAACTTTTGGCA 1152
Db |||||
QY 536 GGAAGCGGAATTTGGCCGAGTGTCTGGAACCAAGGATCTTTACCTTGGAACTTTGCTA 595
Db |||||
QY 1153 CACAAGAAAGCTGCAGAGGCATACGACATTCGCGCCCATCAATTCAGAGGATTAACCG 1212
Db |||||
QY 596 CTGAGGAGGAAGCAGCGGAGCGTGTGACATTCAGCCATTAAGATTCAGGGAGTGAATG 655
Db |||||

1213 CAGTGACTTAACCTCGACATGACAGATACAGGTTAAAGCATCTCTCGAAGCCCTAGTCT 1272
Db |||||
656 CAGTGACCAACTTCGAGATGAATCGTTATGATGTGAGGCGCATCATGCAAG---TTCCC 712
QY 1273 TTCTATTGTTGATCGCGCGCAAAACGCTCTCAAG 1304
Db 713 TCCAGTTGTTGGAACCGCTAGAGCTTTGAAG 744

RESULT 11
AW780688
LOCUS
DEFINITION
AW780688 500 bp mRNA linear EST 03-DEC-2001
sl75e07.y1 Gm-cl027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl027-7165 5' similar to TR:Q41832 Q41832 OPEN READING FRAME. ;
mRNA sequence.
AW780688
AW780688
AW780688.1 GI:7795291
EST.
Glycine max (soybean)
Glycine max
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE
1. (bases 1 to 500)
Shoemaker,R., Keim,P., Vodkin,L., Expelding,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Pearson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 412.
LOCATION/Qualifiers
1. 500
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl027-7165"
/tissue_type="cotyledons of 3- and 7-day-old Williams
seedlings"
/lab_host="DH10B"
/clone_lib="Gm-cl027"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from cotyledons of 3- and 7-day-old Williams seedlings
which were propagated on paper towels with distilled
water. The cotyledons were flash-frozen in liquid
nitrogen, then lyophilized for 72 hours. Unequal amounts
of mRNA was used for cDNA synthesis. Stratagene's cDNA
Synthesis Kit (catalog number 200401) was used to
synthesize the cDNA. First- stranded synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An anchor
nucleotide (V=A, C, or G) was added to the 3' end of the
primer [GAGAGAGAGAGAGAGACTAGTCTCGAG(T)18] to anchor
the primer at the 5' end of the poly(A) tract. After
second- strand synthesis, the cDNA ends were filled in
with cloned Pfu DNA, ligated to EcoRI adapters and
subsequently phosphorylated. The XhoI site within the

```

first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GIBCOBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript (tm) II XR Predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 97% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=30). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

ORIGIN

Query Match 16.7%; Score 335.4; DB 10; Length 500;
Best Local Similarity 79.7%; Pred. No. 6.7e-57;
Matches 396; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

736 AGAGGACGCTCTATATACCGCGGTGTGTACAAAGCATCGTGGACAGAGATATAGGAC 795
Db AGCGCACCTCCATCTACCGCGCGGTCCACCGACATAGATGGACGGAGATACGAAGCTC 63

796 ATTATGGGATATAGTTTAAAGAGAGCGCCAAAGCGGCAAGGACAGAGTTTATT 855
Db ATCTATGGGACATATAGTTTGAAGAGAGGCGCAAGAGGAGAGAGAGTTTACC 123

856 TGGGAGCTTATGACAAAGAGAAAGAGCAGCTAGGCTTTATGATTAGCGCACTCAAGT 915
Db TGGTGCTTATGACAGGAGATAGGCGAGCGGCTTACGATCTCGCAGCTCTCAAGT 183

916 ATTGGGGAACCCACTACTACTTATCTCCCATGAGCGAATATGAAAGAGGTAGAG 975
Db ACTGGGCTCAACTACCAACCACTTTCCCACTTATGAGAGGAGAACTGGAGG 243

976 AGATGAAGACATGACAAAGGCAAGATGTGTGCTCTGCGCAGGAAAGTAGTGGTT 1035
Db AGATGAAGACATGACAAAGGCAAGATGTGTGCTCTGCGCAGGAAAGTAGTGGTT 303

1036 TCTCTCGTGTGCATCGATTTATCGTGGAGTAAAGACATCAACACATGAAAGATGC 1095
Db TCTCTAGGGGGGCTCTATATACAGAGGAGTACGAGACACCAACGATGCGCATGC 363

1096 AAGCTAGGATAGAAAGTCCGCGTAAAGAGACCTTACTTGGGAACTTTTGGCACAC 1155
Db AGCGGAGATAGGACAGATGTCCGGAACAAAGACCTTACCTTGGAACTTTCAGACCC 423

1156 AAGAAGAAGCTGCAGAGGATACGACATTCGGGCCATCAAAATTCAGAGATTAAACGCG 1215
Db AAGAAGAAGCTGCAGAGGCTATGACATTCGCTCTATCAAAATTCAGAGATTAAATGCAG 483

1216 TGACTAACTTCGACATG 1232
Db TCACAACTTTGACATG 500

RESULT 12

BM307997
LOCUS 593 bp mRNA linear EST 02-JAN-2002
DEFINITION sak4e09.y1 Gm-c1036 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1036-5586 5' similar to TR:Q9SYC2 Q9SYC2 FL1M15.6 PROTEIN. 1;
mRNA sequence.
ACCESSION BM307997
VERSION BM307997.1 GI:18039703
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 593)
AUTHORS Shoemaker,R., Keim,P., Vodkin,J., Erpelding,J., Coryell,V.,

Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 427.
Location/Qualifiers
1. 593
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1036-5586"
/tissue_type="somatic embryos cultured on MSD 20"
/lab_host="DH10B"
/clone_lib="Gm-c1036"
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linker adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

FEATURES

source

ORIGIN

Query Match 16.6%; Score 335.2; DB 12; Length 593;
Best Local Similarity 77.5%; Pred. No. 7.1e-57;
Matches 406; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

681 ACTATTGATAGTAGCGTTGAAGCTACACCGAAGAAACTATTGAGAGTTTGGACAGAG 740
Db ACTACTACTACTACTGTGAGCTGCACCTAGAAGACATTGGATACATTCCGGCAGAGA 125

741 ACGTCTATATACCGCGGTGTACAGGCATCGTGGACAGGAATATATAGGCACTTTA 800
Db ACATCCATATATCGTGGAGTAAACGCGACATAGTGGACTGGAAGGTATGAAGCTCACCT 185

801 TGGGATATATAGTGTAAAGAGAGAGGCCAAACCGCGCAAGAAAGCAAGTTTATTGGGA 860
Db TGGGATATATAGTGTAGAGGAGGAGGCAATCAAGAAAGACGCCAAGTTTATTGGGT 245

861 GGTATGACAAAGAGAAAAAGCGCTAGGGCTTATGATTAGCCGCACTCAAGTATTGG 920
Db GGATATGATAAAGAAAAAGACGCTAGGTCTTATGATTAGCTGCACTGAAGTACTGG 305

921 GGACCCACTACTACTTAAGTCCCATGACCGAATATGAAAGAGGTAGAGAGATG 980
Db GGACATCCACCATCACCACCTTCCCAATTAGTACTATGAGAGGAATGATGATGATG 365

981 AAGCATCATGACAAAGCAAGAGTATTTGCTCTCACTGCCAGGAAAGTAGTGGTTCTCT 1040
Db AAACATCATGACGCGCAAGAAATTTGCTGCTCATTAGAGGAAAAAGCAGTGGTTCTCC 425

```

QY 1041 CGTGTGTCATCGATTATTCGTGGAGTAAACAGACATCACCAACATGGAAGATGCGAGCT 1100
Db 426 AGGGTGCATCAATGATATCGTGGAGTATCAAGGCATCACCAACAGGAATGCGAAGCA 485
QY 1101 AGGATAGGAAGTGGCGGTAAACAAAGACCTCTACTTGGGAATTTTGGCACACAGAA 1160
Db 486 AGAATTGCGAGAGTTCGAGGAACAAAGATCTTTACTTGGGAATTTTCACTACTGAGAA 545
QY 1161 GAAGCTGCAGAGGATACGACATTCGCGCCATCAAAATTCAGAG 1204
Db 546 GAGCTGCTGAAGCATACGACATAGCTGCGATAAAGTTTCAGAG 589

RESULT 13
BZ044547
LOCUS
DEFINITION BZ044547 717 bp DNA linear GSS 09-OCT-2002
1k152h10.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
sequence.
ACCESSION BZ044547
VERSION BZ044547.1 GI:23638096
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 717)
AUTHORS Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
Nash,W., Rabinowicz,P.D. and Wilson,R.K.
TITLE Whole genome shotgun reads from Brassica oleracea
JOURNAL Unpublished (2002)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: 1k152 row: h column: 10
Seq primer: -21uppoT forward
Class: shotgun
High quality sequence start: 16
High quality sequence stop: 551.

FEATURES
source
Location/Qualifiers
1..717
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/notes="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea T01000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN
Query Match 16.6%; Score 333.6; DB 28; Length 717;
Best Local Similarity 97.4%; Pred. No. 1.4e-56;
Matches 339; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 GTTCATCTCTCTTTTAAAGACCAAAACCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
Db 313 GTTCATCTCTCTTTTAAAGACCAAAACCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 372
QY 61 ACTAAGTTCTCTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTCTCTGAACCTCAA 120
Db 373 ACTAAGTTCTCTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTCTCTGAACCTCAA 432
QY 121 TGAATATACTGGTAGGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db 433 TGAATATACTGGTAGGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 492
QY 181 ACGTCTACTCTTCCACCACCAACCGTCGTAGATGTCGCGGAGAGTACTGTTCAGATC 240

```

```

Db 493 AGCTCTACTCTTCCACCACCAACCGCGTAGATGTCGCGGAGACTACTGTTCAGATC 552
QY 241 CGACCGCTGCTCCGATGAGTCTTCAGGCATCCAAACATCTTTCTCTCTCTCTCTCTCTCTCT 300
Db 553 CGACCGCTGCTCCGATGAAATCTTTAGGCATCCAAACATCTTTCTCTCTCTCTCTCTCTCT 612
QY 301 TCGTCTGTGATGCTTTTACCAGAGACAAATAGTCACTCCCGAGATT 348
Db 613 TCGTCTGTGATGCTTTTACCAGAGACAAATAGTCACTCCCGAGATT 660

RESULT 14
BZ496117
LOCUS
DEFINITION BZ496117 686 bp DNA linear GSS 16-DEC-2002
BONPX49TR BO_1.6_2_KB_tot Brassica oleracea genomic clone BONPX49,
genomic survey sequence.
ACCESSION BZ496117
VERSION BZ496117.1 GI:27008347
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 686)
AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BONPX49TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..686
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone_lib="BONPX49"
/notes="Vector: pBONPX49; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pBONPX49 using BstXI linkers"

ORIGIN
Query Match 16.5%; Score 332; DB 28; Length 686;
Best Local Similarity 97.1%; Pred. No. 3e-56;
Matches 338; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GTTCATCTCTCTTTTAAAGACCAAAACCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
Db 335 GTTCATCTCTCTTTTAAAGACCAAAACCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 394
QY 61 ACTAAGTTCTCTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTCTCTGAACCTCAA 120
Db 395 ACTAAGTTCTCTCTTTTACCTTTTACCAAGAACTCGTTAGATCACTCTCTGAACCTCAA 454
QY 121 TGAATATACTGGTAGGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Db 455 TGAATATACTGGTAGGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 514
QY 181 ACGTCTACTCTTCCACCACCAACCGTCGTAGATGTCGCGGAGAGTACTGTTCAGATC 240
Db 515 ACGTCTACTCTTCCACCACCAACCGTCGTAGATGTCGCGGAGAGTACTGTTCAGATC 574
QY 241 CGACCGCTGCTCCGATGAGTCTTTCAGCCATCCAAACATCTTTCTCTCTCTCTCTCTCTCTCT 300

```

Db 575 CGACGGCTCGCTCCGATGAATCTTTAGCCATCCAAACATCGTTTCCCTTCCTCCCTTTGGTG 634

QY 301 TCGTGTGATGCTTTACACGAGACAACATAGTCACTCCCGAGATT 348

Db 635 TCGTCTCGATGCTTTACACGAGACAACATAGTCACTCCCGAGATT 682

RESULT 15
CD475982
LOCUS

DEFINITION
nadh03-15msl-h07 Nad03 Nuphar advena cDNA clone nad03-15msl-h07 5',
mRNA sequence.

ACCESSION
CD475982.1 GI:31397150

VERSION
CD475982.1

KEYWORDS
EST.

SOURCE
Nuphar advena

ORGANISM
Nuphar advena

REFERENCE
AUTHORS
dePamphilis, C., Carlson, J., Ma, H., Frohlich, M., Tanksley, S.,
Leebens-Mack, J., Field, D., Arrington, J., Zahn, D., Kong, H.,
Druckemiller, M., Landherr, L., Hu, Y., Ilut, D., Wall, K.,
Plock, S., Chiorean, S., Albert, V., Doyle, J., Miller, W.,
Oppenheimer, D., Soltis, D., Soltis, P. and Theissen, G.
Generation of ESTs from early flower buds of Nuphar advena
Unpublished (2002)

TITLE
JOURNAL
COMMENT
Contact: Claude dePamphilis or James Leebens-Mack
Mueiller Laboratory
Penn State University
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 863 8413
Fax: 814 865 9131
Email: cwd3@psu.edu or jhl10@psu.edu
The sequence provided is trimmed of vector and low quality regions.
Full sequence and original trace file are available from the Plant
Genome Network website (<http://pgn.cornell.edu>)
Plate: nad03-15msl row: h column: 07
Seq primer: M13P.

FEATURES
source
1..545
/organism="Nuphar advena"
/mol_type="mRNA"
/db_xref="taxon:77108"
/clone="nad03-15msl-h07"
/tissue_type="flower buds"
/dev_stage="c= 2.5mm"
/lab_host="SOLR"
/clone_lib="Nad03"
/notes="Vector: pBluescript SK (-/-); Site 1: EcoRI;
Site 2: XhoI; Only floral buds with diameter of 2.5 mm of
less were used for RNA isolation. This is a directionally
cloned, non-normalized library. Avg. insert length: 1134;
Primers: M13P and M13R; Antibiotic: 50 ug/ml Ampicillin;
Primary titer: 2E6 pfu total; Amplified titer: 3.2E10
pfu/ml; Mass Excised Titer: SE10 total; This library has
been generated by the Floral Genome Project (FGP). The
Floral Genome Project is funded by NSF's Plant Genome
Research Program (DBI-0115684). More information about the
project can be obtained at <http://fgp.bio.psu.edu>

ORIGIN
Query Match 16.1%; Score 323.6; DB 14; Length 545;
Best Local Similarity 75.1%; Pred. No. 1.5e-54;
Matches 404; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 703 CTACACCGAGAAACATTTAGAGTTTTCGACAGAGGACGCTATATACCGCGGTGTTA 762

Db 5 CTATGCATAGGAGTCCATTGACACCTTTGGACAGAGAACATCTTCAGTATAGAGGGGTTA 64

QY 763 CAAGGCATCGGTGGCAGGAGAGATATGAGGCACATTTATGGGATATAGTTGTAAAGAG 822

Db 65 CAAGACACAGATGGACTGTAGATATGAAGCCCATCTTTGGGACATAGCTGTAGAGG 124

QY 823 AAGGCCAAACCGCCAAAGGAAGCAAGTTTATTTGGGAGGTTATGACAAAGAAAGAAAG 882

Db 125 AAGGGCAAAACCAAGAAAGGAAGCAAGTCTATCTGGGTGGCTATGACATGGGAAGAAAG 184

QY 883 CAGCTAGGGCTTATGATTTAGCCGCACTCAAGTATTGGGGAAACCCACCTACTACTAACT 942

Db 185 CAGCTAGGGCTTATGACTTTAGCAGCACTCAATATTTGGGGGCTGTCTACTCACAAT 244

QY 943 TCCCATGAGCGCAATATGAAGAGAGGTGAAGAGATGAAGCAATGACAGCAAGAGT 1002

Db 245 TTCCATTGAGCAATTACCAGGAAGAACTTGAAGGGATGAAGAAACATGACAGCAAGAGT 304

QY 1003 ATGTTGCTCAGTCCGCGAGGAAAGTAGTGTCTCTCGTGGTGCATCGATTATCGTG 1062

Db 305 ATGTTGCTCAGTTAAGGAGGAAAGTAGTGGGTTTCAAGAGCGGATCTATGTATAGAG 364

QY 1063 GAGTAAACAAGACATCACCAACATGGAAGATGGCAAGCTAGGATAGGAAGAGTCCCGGTA 1122

Db 365 GTGTAACAAGGCATCATCAACATGGGAGGTGGCAAGCTAGGATTGGTAGAGTAGCTGAA 424

QY 1123 ACAAGACCTCTACTTTGGGAACTTTTGGCACACAAGAGAGCTGCAGAGGCATACGACA 1182

Db 425 ACAAGACCTTTTATCTTTGGGACCTTTAGTACTCAGGAGGAGGAGCTGAAAGCATACGCG 484

QY 1183 TTGGGGCCATCAAAATTCAGAGGATTACCCGAGTGAATCTTCGACATGAACAGATA 1240

Db 485 TTGCTGCAATCAAAATTCGAGGGTTGATGCTGTGACCAACTTTGACATAACAGATA 542

Search completed: July 19, 2004, 10:20:06
Job time : 5297 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2004, 05:43:48 ; Search time 156 Seconds
(without alignments)
7164.558 Million cell updates/sec

Title: US-09-980-364-1

Perfect score: 2014

Sequence: 1 gttcatctcttcttcttaag.....gttgataaatttcagact 2014

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/prodata/2/ina/5A-COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B-COMB.seq.*
3: /cgn2_6/prodata/2/ina/5A-COMB.seq.*
4: /cgn2_6/prodata/2/ina/5B-COMB.seq.*
5: /cgn2_6/prodata/2/ina/5A-COMB.seq.*
6: /cgn2_6/prodata/2/ina/5B-COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	330.6	16.4	2148	4	US-09-227-421-1
2	330.6	16.4	2148	4	US-09-479-855-1
3	324.8	16.1	1738	4	US-09-227-421-4
4	324.8	16.1	1738	4	US-09-479-855-4
5	129	6.4	1680	2	US-08-700-152A-3
6	79.8	4.0	301	4	US-09-313-294A-6727
7	67.2	3.3	7218	1	US-08-232-463-14
8	55.2	2.7	234	4	US-09-313-294A-6087
9	55	2.7	7218	1	US-08-232-463-14
10	44.6	2.2	5562	4	US-10-204-708-63
11	44.4	2.2	1024	4	US-09-328-475C-50
12	43.4	2.2	8607	4	US-10-204-708-72
13	43.2	2.1	658	3	US-08-998-416-595
14	43.2	2.1	6040	4	US-10-204-708-69
15	42.8	2.1	11721	4	US-09-026-039-3
16	42.8	2.1	11721	4	US-09-026-039-3
17	42.6	2.1	6368	4	US-10-204-708-68
18	42.4	2.1	396	4	US-09-640-173-53
19	42.4	2.1	396	4	US-09-713-550-53
20	42.4	2.1	731	1	US-08-451-405A-2
21	42.2	2.1	240	1	US-08-628-417-6
22	42.2	2.1	5535	4	US-10-204-708-18
23	42	2.1	5152	4	US-10-204-708-47
24	41.8	2.1	674	4	US-09-620-405B-465
25	41.8	2.1	674	4	US-09-433-826B-465
26	41.8	2.1	674	4	US-09-604-287A-465
27	41.8	2.1	674	4	US-09-834-759-465

28	41.8	2.1	8537	4	US-10-204-708-41	Sequence 41, Appl
29	41.4	2.1	1117	3	US-09-247-373B-33	Sequence 33, Appl
30	41.4	2.1	2447	2	US-09-014-969-14	Sequence 14, Appl
31	41.4	2.1	2634	4	US-09-463-238-3	Sequence 3, Appl
32	41.4	2.1	6659	4	US-10-204-708-6	Sequence 6, Appl
33	41.2	2.0	1378	4	US-09-149-476-208	Sequence 208, Appl
34	41.2	2.0	6583	4	US-10-204-708-26	Sequence 26, Appl
35	40.8	2.0	174	4	US-09-302-161B-21	Sequence 21, Appl
36	40.8	2.0	1048	4	US-09-489-847-38	Sequence 38, Appl
37	40.8	2.0	1405	4	US-09-202-161B-6	Sequence 6, Appl
38	40.6	2.0	2323	4	US-09-149-476-24	Sequence 24, Appl
39	40.6	2.0	7286	3	US-09-331-581-3	Sequence 3, Appl
40	40.6	2.0	7938	3	US-09-331-581-14	Sequence 14, Appl
41	40.6	2.0	8961	4	US-10-204-708-80	Sequence 80, Appl
42	40.6	2.0	19124	2	US-08-487-826B-13	Sequence 13, Appl
43	40.4	2.0	1558	1	US-08-467-607-2	Sequence 2, Appl
44	40.4	2.0	1558	2	US-08-469-362-2	Sequence 2, Appl
45	40.4	2.0	1558	2	US-08-850-392-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-227-421-1
; Sequence 1, Application US/09227421
; Patent No. 6559357
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Mizukami, Yukiko
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility
; FILE REFERENCE: 023070-090700PC
; CURRENT APPLICATION NUMBER: US/09/227,421
; PRIOR FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: US 09/227,421
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2148
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (269)..(1936)
; OTHER INFORMATION: AINTEGUMENTA (ANT)
; NAME/KEY: misc_feature
; LOCATION: (1109)..(1339)
; OTHER INFORMATION: encodes first AP2 domain
; NAME/KEY: misc_feature
; LOCATION: (1340)..(1414)
; OTHER INFORMATION: encodes linker region
; NAME/KEY: misc_feature
; LOCATION: (1415)..(1621)
; OTHER INFORMATION: encodes second AP2 domain
US-09-227-421-1

Query Match 16.4%; Score 330.6; DB 4; Length 2148;
Best Local Similarity 75.4%; Pred. No. 9.7e-85;
Matches 411; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY	714	AAACTATTGAGATTTTGGACAGACGCTCTATATACCGCGTGTACAGGCAATCGG	773
Db	1082	AAATCTATCGTACTTTTGGACACGAACTTCTCAATACCGAGCGTTTACAGACATAGA	1141
QY	774	TGGACAGGAAGATGAGGCACATTTATGGGATATAGTTTAAAGAGAGGCAACG	833
Db	1142	TGGATCTGTAGATATGAGCTCATCTATGGACATAGTTTCAAGAGGAGGTACAGT	1201
QY	834	CCCAAGGAAGACAAAGTTTATTTGGAGGTTATGCAAGAGAAAGAGAGCTTAGGCT	893

1202	AGAAAGGAAGACAAGTTTATCTGGGAGGTTATGATATGGAGGAGAAAGCTGCTCGAGCA	1261
894	TATGATTATAGCCGCACTCAAGTAGTTGGGAACCAACCACTACTACTAACTTCCCATGAGC	953
1262	TATGATCTTGCTGCACCTCAAGTACTGGGTCCCTCTACTCACACCAATTTCTCTGGGAG	1321
954	GAATATGAAGAAGAGGTAGAAGATGAAGCACATGACAAGGCAAGAGTATGTTGCTCA	1013
1322	AATTATCAGAAAGAGATTGAAGACATGAAGACATGACTAGACAAGAATATGTCACAT	1381
1014	CTGCGCAGAGAAAGTAGTGGTTTTCTCTCGTGGTGCAATCGATTTATCGTGGAGTAACAAGA	1073
1382	TTGAGAGAAGAGCAGTAGTGGTTTTCTCTAGGGGTGCTTCATCTATAGAGGTACAAGA	1441
1074	CATCACCAACATGGAAAGATGGCAAGCTAGGATAGGAAGAGTCGCCGGTAACAAGACCTC	1133
1442	CATCACCAAGCATGGAAGGTGGCAAGCAGCGATTGGTAGAGTCGCTGAAACAAGATCTC	1501
1134	TACTTGGGAACCTTTTGGCACAAGAAGAAGCTGCAGAGGCATACGACATTCGGCGCATC	1193
1502	TACCTTGGAACTTTTGGAAACCAAGAAGAAGCTGCAGAAGCTTACGATGTAGCAGCAATT	1561
1194	AAATTCCAGAGGATTAACCGCAGTGACTTACTTCGACATGAACAAGATACAACTGTAAGA	1253
1562	AGTTCCGTGGCACAATGCTGTGACTTACTTTGATATCACGAGGTACGATGTTGCTG	1621
1254	ATCCT	1258
1622	ATCAT	1626

```

RESULT 2
US-09-479-855-1
; Sequence 1, Application US/09479855
; Patent No. 6639128
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Mizukami, Yukiko
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility
; TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants
; FILE REFERENCE: 023070-090720US
; CURRENT APPLICATION NUMBER: US/09/479,855
; CURRENT FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2148
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: AINTEGUMENTA (ANT) cDNA
; NAME/KEY: CDS
; LOCATION: (269)..(1936)
; OTHER INFORMATION: AINTEGUMENTA (ANT)
US-09-479-855-1

```

Qy	894	TATGATTTACCGC	ACTCAAGTATT	GGGGAACCA	CACTACTACT	TACTTCCCAATGAGC	953
Db	1262	TATGATCTTGC	TGCATCACTCAAGTACT	GGGGTCCCTCTACT	CAACCAATTTCTCTCGGGAG	1321	
Qy	954	GAATATGAAAAGAG	GTTAGACAGATG	AGCACATGCA	CAAGGCRAGAGT	GTGTTGCCCTCA	1013
Db	1322	AATTATCAGAAA	GAGATTGAAGACAT	GAAAGACATG	ACTAGACAAGAAT	ATGTTGCACAT	1381
Qy	1014	CTGCGCAGGAAAA	AGTAGTGGTTTCT	CGTGGTGCAT	CGATTTATCGTGGAGTAA	CAAGA	1073
Db	1382	TTGAGNAGAA	AGCGAGTGGTTTCT	CTAGGGGTGCTTCCA	CTATAGAGGAGTCA	CAAGA	1441
Qy	1074	CATCACCAACAT	GGAAGATGGCAAGCT	AGGATGGAAGAGT	GC CGGTAA	CAAGA	1133
Db	1442	CATCACCA	GCATGGAAGTGGCAAGC	CGGATTGGTAGAGT	CGCTGGAACAAAGATCTC	1501	
Qy	1134	TACTTTGGG	AACTTTTGGCACAAAG	AGCTGCGAGAGGC	ATACGATTTGCGGCCATC	1193	
Db	1502	TACCTTGG	AACTTTTTGGAACCCAAAG	AAAGCTGCAGAAAGCTT	ACGATGTAGCAGCAAT	1561	
Qy	1194	AAATTCAGAGG	ATTAAACCGCAGTGACT	TAACTTCGACATGAA	CAGATACAAAGCTTAAAGCA	1253	
Db	1562	AGGTTCCGTGG	CACAAATGCTGTGACT	TAACTTTGATATCA	CGAGTACGATGTTGATCGT	1621	
Qy	1254	ATCCT	1258				
Db	1622	ATCAT	1626				

```

RESULT 3
US-09-227-421-4
; Sequence 4, Application US/09227421
; Patent No. 6559357
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Mizukami, Yukiko
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility
; TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants
; FILE REFERENCE: 023070-090700PC
; CURRENT APPLICATION NUMBER: US/09/227,421
; CURRENT FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: US 09/227,421
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1738
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1647)
; OTHER INFORMATION: BANT, ANT ortholog from Brassica napus (Carolia)
US-09-227-421-4

```

Query Match	16.4%;	Score 330.6;	DB 4;	Length 2148;
Best Local Similarity	75.4%;	Pred. No. 9.7e-85;		
Matches 411;	Conservative 0;	Mismatches 134;	Indels 0;	Gaps 0;
714 AAAACTATTGACAGTTTTGGACAGAGGACGTCATATACCGCGGTGTTAAAGGCGATCGG	773			
1082 AAATCTATCGATACTTTTGGACAACGAACTTCTCAATACCGAGGCGTTACAAGACATAGA	1141			
774 TGGACAGGAAGATATAGGCCACATTTATGGGTAATAGTTGTAAGAGAGGCCCAACG	833			
1142 TGGACTGCTAGATATGAAGCTCATCTATGGGCAATAGTTTCAAGAAAGGAGGTCACAGT	1201			
834 CGCAAGGAGACAAGTTTATTTGGGAGGTTATGACAAGAGAAAAAGCAGCTAGGGCT	893			
1202 AGAAAGGGAACAAGTTTATCTGGGAGGTTATGATATGGAGAGAAAGCTGCTCGAGCA	1261			

Query Match	16.1%;	Score 324.8;	DB 4;	Length 1738;
Best Local Similarity	73.8%;	Pred. No. 4e-83;		
Matches 413;	Conservative 0;	Mismatches 147;	Indels 0;	Gaps 0;
QY	714	AAAACTATTGAGAGTTTGGACAGAGGACGCTATATACCGCGGTGTGTACAAAGCATCGG	773	
DB	790	AAATCTATTGTATCTTTTGGACAGAACTTCGGCAATACCGAGCGTTTCAAGACATAGA	849	
QY	774	TGGACAGGAAGATATGAGCGCATTTATGGGATATATAGTTCTAAAGAGAAAGGCCAAACG	833	
DB	850	TGGACTCGGTATGAAGCTCATCTATGGGACAAATAGTTTCAAGAAAGGAAGGTCATAGC	909	
QY	834	CGCAAGGAAGACAAGTTTATTTCGGAGGTTATGACAAAGAAAGAAAGCAGCTAGGGCT	893	
DB	910	AGAAAGGAAGACAAGTTTATTCGGGGGGTTTATGATATGGAGGAGAAGCTGCTCGAGCA	969	
QY	894	TATGATTTTAGCGGCACCTCAAGTATTTGGGGAAACCACTACTACTACTTCCCATCAGC	953	

Best Local Similarity 57.4%; Pred. No. 7.5e-27;
Matches 253; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

QY 840 GGAAGACAAAGTTATTTGGGAGGTATGACAAAGAAAAGCAGTAGGGCTTATGAT 899
DB 643 GGGAAACAAGTTTACTTAGGTGGATTTGACACTGCTCATGAGCAGCTCGAGCATATGAT 702
QY 900 TTAGCCGCACTCAAGTATTTGGGAAACACCACTACTACTACTTCCCATCAGCGATAT 959
DB 703 AGAGCTGCTATTAATTCCTGGAGTAGAACGGGATATCAATTTCAACATCGAGATAT 762
QY 960 GAAAAGAGGTAGAGAGATCAAGCACAATGACAAAGGCAAGAGTATGTTGCCCTCAGTCGC 1019
DB 763 GATGATGACTTTGAAACAGATGACTAATTTAAACCAAGGAAGTTCGTACACGTAATTCGC 822
QY 1020 AGGAAAGTAGTGTCTCTCGTGGTGCATCGATTTATCGTGAGTAAACAAGACATCAC 1079
DB 823 CGAACAAGCAGGCTTCCCTCGAGGAAGTTTCGAAGTATAGAGGTGCATTTGCAAT--- 879
QY 1080 CAACATGGAAGATGCAAGCTAGGATAGGAAGAGTCCGCGTAAACAAGACCTCTACTTG 1139
DB 880 AAGTGTGCTGCTGGGAAGCTCGTATGGTCAATTTCTAGGCAAAAGTATGTTATTG 939
QY 1140 GGAAGTTTGGACACAAAGAAAGAGCTGCAGAGGCATACGACATTTGGGCGCATCAATTC 1199
DB 940 GGTGTTTCGACACCGAGGTCGAAGCTGTAGAGCTTACGATAAAGCTGCAATCAAAATGT 999
QY 1200 AGAGGATTAACCGCAGTGAATCTTCGACATGAACAGATACAACTGTTAAAGCAATCCTC 1259
DB 1000 AACGCAAGAGCGCGTGACCAATTTTGATCCGAGTATTTACGATGAGGAACTCAATGCC 1059
QY 1260 GAAAGCCCTAGTCTCTCTATT 1280
DB 1060 GAGTCATCAGGGAATCCTACT 1080

RESULT 6

US-09-313-294A-6727
; Sequence 6727, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: EL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6727
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700352302H1
; NAME/KEY: unsure
; LOCATION: 297
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-6727

Query Match 4.0%; Score 79.8; DB 4; Length 301;
Best Local Similarity 60.3%; Pred. No. 3.9e-13;
Matches 149; Conservative 0; Mismatches 97; Indels 1; Gaps 1;

QY 848 AGTTTATTTGGGAGGTATGACAAAGAAAAGCAGTAGGGCTTATGATTTAGCCGC 907
DB 50 AGTCTACTGGGGCGTATGACACGAGGAGGAGCGTCTTCGTGCGCTATGACCTCGAGC 109
QY 908 ACTCAAGTATTTGGGAAACCACTACTACTACTTCCCATCAGGCAATATGAAAAGA 967
DB 110 TCTCAAGTACTGGGCTCTGAGACTCTGCTCAACTTCCCTGTGAGGATTAATCCAGCA 169

QY 968 GGTAGAGAGATGAAGCAGCATGACAAGGCAAGAGTATGTTGCTTACTGCGGAGAAAAG 1027
DB 170 GATGCGGAGATGGAGCGCGTGTCCGGGAGGAGTACCTGGCTTCCCTCCGCGAGAG 229
QY 1028 TAGTGGTTTCTCTCGTGGTGCATCGATTTATCGTGGAGTAAACAAGACATCAACACATGG 1087
DB 230 CAGCGGCTTCTCAGGGCGCTCTCCAAGTACAGAGCGCTCGC-AGGCATCACCACACGG 288
QY 1088 AAGATGG 1094
DB 289 GAGTGGG 295

RESULT 7

US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 3.3%; Score 67.2; DB 1; Length 7218;
Best Local Similarity 7.5%; Pred. No. 1e-08; 175; Indels 0; Gaps 0;
Matches 33; Conservative 232; Mismatches 175; Indels 0; Gaps 0;

QY 596 AAAAGGCTGTCCTCTCAATGAATCATCTTCTTGTGATTAACAACACGACGAA 655
DB 1504 AAAGCGCATGTAGGCATCACTGTATTACCTATCTATGCAAGTAGTTAAAGAGATAGAAG 1445
QY 656 TAACAACGTTTTCGCCAAGGGAAGACTATTGTGATAGCTTGAAGCTACACCGAAGAA 715

[illegible]

```

RESULT 8
US-09-313-294A-6087
; Sequence 6087, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6087
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700351215H1
US-09-313-294A-6087

```

	Query Match	2.7%	Score 55.2	DB 4	Length 294	
	Best Local Similarity	75.0%	Pred. No. 4.5e-06			
	Matches 69	Conservative 0	Mismatches 23	Indels 0	Gaps 0	
Qy	868	ACAAGAAGAAAAAGCAGCTAGGCGCTTATGATTTAGCGGCACTCAAGTATTGGGGAACCA	927			
Db	3	ATAAGAAGAGAAAGCTACCAAGGCTTACAAATTTAGCTGCTCTCAAGTACTGGGGAACCA	62			
Qy	928	CCACTACTACTAACTTCCCGCATGACGCAATAT	959			
Db	63	CAACGACTTACAAAGTTTTCCAGTATTTGCATAT	94			

RESULT 9
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52

CORESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Hardner
 STREET: 1800 Diagonal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22313-0299
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/232,463
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/935,313
 FILING DATE:
 APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)836-9300
 TELEFAX: (703)683-4109
 TELEX: 899149
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 CLONE: PTZgpt-F1s
 US-08-232-463-14

Query Match 2.7%; Score 55; DB 1; Length 7218;
Best Local Similarity 2.9%; Pred. No. 3.2e-05;
Matches 10; Conservative 203; Mismatches 128; Indels

Qy	2	TTCAATCTCTTCTTTAAGACCAAACCTTTTCTCCTCTCTTCATGCATGAACCTAA	61
Dd	1114	YVVVYYY	1173
Qy	62	CTAAGTCTCTCTTTTACTTTTACCAGAATCGTTAGATCACTCTCTGAACCTAAT	121
Dd	1174	YVVVYYY	1233
Qy	122	GAATAATACTGTTAGGCTTTTCTCTCTCTATTGAACAAAATCACCATCGTAAGGA	181
Dd	1234	YVVVYYY	1293
Qy	182	CGTCTACTTCCACCACCAACCGTCGTAGATGTGGCGGAGAGTACTGTACGATCC	241
Dd	1294	YVVVYYY	1353
Qy	242	GACCGTGCTCCGATGAGTCTTCAGCCATCCAAACATCGTTCTCTCCCTTGTGTGT	301
Dd	1354	YVVVYYY	1413
Qy	302	CGTCGTGATGCTTCCAGACAGACAATAGTCACTCCC	342
Dd	1414	YVVVYYYYYYYYYYYYYYYYYGTACCAAAATCTCTAATC	1454

RESULT 10
US-10-204-708-63
; sequence 63, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:

```
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 63
; LENGTH: 5562
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-63
```

```
Query Match      2.2%; Score 44.6; DB 4; Length 5562;
Best Local Similarity 58.8%; Pred. No. 0.027;
Matches 77; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 1878 TTTAGTATATGCGTTGTCGCTGCGTGAACAGTGTGATCTTTCATTGATGTTTTT 1937
Db 1671 TTTTATATATATTTATTTTGTGTTTTTATTTATGATGTTTTTTTTTTTTCG 1730

QY 1938 CTTTCTCTTTCTTTTCTGTTGTTAAATCTTAAAGACTTATTTAGTTCCATTGTT 1997
Db 1731 TTTTATGATGTTTTTTTACGTTTTCGTTTTGAGAAATATGATGTTTTTTTAGTA 1790

QY 1998 GGATAAAATTTT 2008
Db 1791 TGGTAGTTTTT 1801
```

```
RESULT 11
US-09-328-475C-50/c
; Sequence 50, Application US/09328475C
; Patent No. 6476207
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jimmy
; APPLICANT: Astel, Jon H.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ford, Donna M.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Steinmann, Kathleen E.
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
; TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER
; FILE REFERENCE: 1532.002/200130.463
; CURRENT APPLICATION NUMBER: US/09/328,475C
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 50
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1024)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-475C-50
```

```
Query Match      2.2%; Score 44.4; DB 4; Length 1024;
Best Local Similarity 45.2%; Pred. No. 0.012;
Matches 90; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 1810 AAGGTGGTGGAGAAGTTGCTCCAACTTTACAGTTTGGAAACGACAATTAGAAAAATAGT 1869
Db 353 AAAAAANNNGNANNAGAGGNCNANCCNNNAANANNCGAAANNNTATANGATNNCCCT 294

QY 1870 TAAAGATCTTTAGTTATATGCGTTCTGTCGCTGGAACAGTGTGATCTTTCATTAT 1929
Db 293 AAAAGATTTTANNNGNANCTTTACCTTNAATTTTTTTTANTNTNTTTTTTTTTTT 234

QY 1930 GTTTTTCTTTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 1989
Db 233 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 174

QY 1990 CATTAGTTGGATAAATTTT 2008
Db 173 TTTTTTTTTTTTTTTTTTTT 155
```

```
RESULT 12
US-10-204-708-72
; Sequence 72, Application US/10204708
; Patent No. 6877731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 72
; LENGTH: 8607
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-72
```

```
Query Match      2.2%; Score 43.4; DB 4; Length 8607;
Best Local Similarity 57.9%; Pred. No. 0.077;
Matches 77; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1876 TCTTTAGTTATATGCGTTGTTGTCGCTGCGTGAACAGTGTGATCTTTCATTGTTTTT 1935
Db 6988 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7047

QY 1936 TCTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1995
Db 7048 TTTTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTTT 7107

QY 1996 TGGATAAATTTT 2008
Db 7108 TTATATATTTTTT 7120
```

```
RESULT 13
US-08-998-416-595
```

```

; Sequence 595, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgin
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebeschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NO. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998/416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 595:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 658 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1408RP
; US-08-998-416-595

Query Match 2.1%; Score 43.2; DB 3; Length 658;
Best Local Similarity 55.3%; Pred. No. 0.02;
Matches 84; Conservative 0; Mismatches 68; Indels

QY 1857 TAGAAAAATAGTAAAGATCTTTAGCTATATGCGTGTGTGCTGCTGTAATTC
DB 493 TAGAGATTGATTTTGTGAACATATATATGCTATTTCTATTCTATTCTATT
QY 1917 ATACTTGTATTATGTTTTTTTCTTTCTCTTTTCTTTTCTTTTCTGGTAATTC
DB 553 TTAATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT
QY 1977 TATTTTGTATTTCCATTAGTTGGATAAATTTT 2008
DB 613 TTAATTTATTTTATTTTATTTTATTTTATTTT 644

RESULT 14
US-10-204-708-69
; Sequence 69, Application US/10204708
; Patent No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander

```

```
/ FILING DATE: 20-JUN-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/700,152
/ FILING DATE: 20-AUG-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bastian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 023070-067230US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 11721 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: -
/ LOCATION: 1..11721
/ OTHER INFORMATION: /note= "Arabidopsis APETALA2 (AP2)"
/ OTHER INFORMATION: Genomic sequence"
US-09-026-039-3
```

```
Query Match          2.1%; Score 42.8; DB 4; Length 11721;
Best Local Similarity 53.9%; Pred. No. 0.14;
Matches 111; Conservative 0; Mismatches 92; Indels 3; Gaps 1;

QY 955 AATATGAAAAAGAGGTAGAGAGATGAAGCACATGCACAAGGCAAGAGTATGTTGCCTCAC 1014
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2035 AATATTGAAATTGAATTTTATAGATGACTAATTTAACCAAGGAAGCTTCGTACCGTAC 2094
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1015 TGGCAGCAAAAGTACTAGTGGTTTCTCTCGTGTGTCATCGATTTATCGTGGAGTAACAAGAC 1074
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2095 TTCGCCGACAAAGCACAGGCTCCCTCGAGGAGTTCGAAAGTATAGAGGTGTCACTTTGC 2154
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1075 ATCAACACATGGAAGATGGCAAGCTAGGATAGGAAGATCGCGCGGTAAACAAAGACCTCT 1134
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2155 ATAAGTG---TGGTCGTTGGGAAGCTCGAATGGGTCAATTCCTTAGGCAAAAGTATAATT 2211
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1135 ACTTGGGAACCTTTGGCACACAGAA 1160
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2212 TCTCTCATTTTATATTCACCTCGAAA 2237
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: July 19, 2004, 10:22:49
Job time : 160 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 19, 2004, 08:51:52 ; Search time 921 Seconds

(without alignments)
10868.347 Million cell updates/sec

Title: US-09-980-364-1

Perfect score: 2014

Sequence: 1 gttcattctcttcttaag.....gttgataaatttcagact 2014

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3190992 seqs, 2439311697 residues

Total number of hits satisfying chosen parameters: 6381984

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA.*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
17: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*
18: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*
19: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	420.6	20.9	2463	13	US-10-183-687-355 Sequence 355, App
2	413.6	20.5	1824	13	US-10-183-687-353 Sequence 353, App
3	413.6	20.5	2588	13	US-10-183-687-351 Sequence 351, App
4	385.8	19.2	1990	13	US-10-183-687-361 Sequence 361, App
5	381	18.9	2088	16	US-10-374-780A-1738 Sequence 1738, App
6	380.8	18.9	2169	13	US-10-424-599-91337 Sequence 91337, A
7	377.4	18.7	1818	13	US-10-225-066A-365 Sequence 365, App
8	377.4	18.7	1818	13	US-10-225-066A-855 Sequence 855, App
9	377.4	18.7	1818	13	US-10-225-066A-911 Sequence 911, App
10	377.4	18.7	1818	16	US-10-374-780A-333 Sequence 333, App
11	364.4	18.1	794	13	US-10-425-114-20762 Sequence 20762, A
12	364.4	18.1	806	16	US-10-374-780A-1742 Sequence 1742, App
13	357.4	17.7	1935	13	US-10-183-687-486 Sequence 486, App
14	357	17.7	1479	13	US-10-425-114-14859 Sequence 14859, A

US-10-424-599-70525 Sequence 70525, A
US-10-183-687-357 Sequence 357, App
US-10-183-687-359 Sequence 359, App
US-10-183-687-490 Sequence 490, App
US-10-424-599-123485 Sequence 123485
US-10-183-687-482 Sequence 482, App
US-10-425-114-19324 Sequence 19324, A
US-10-183-687-347 Sequence 347, App
US-10-374-780A-1763 Sequence 1763, App
US-10-424-599-100930 Sequence 100930, A
US-10-425-114-11685 Sequence 11685, A
US-10-225-066A-1013 Sequence 1013, App
US-10-374-780A-2495 Sequence 2495, App
US-10-059-911-27 Sequence 27, App
US-10-059-911-26 Sequence 26, App
US-10-059-911-38 Sequence 38, App
US-10-059-911-34 Sequence 34, App
US-10-059-911-32 Sequence 32, App
US-10-059-911-30 Sequence 30, App
US-10-183-687-484 Sequence 484, App
US-10-425-114-783 Sequence 783, App
US-10-183-687-488 Sequence 488, App
US-10-059-911-39 Sequence 39, App
US-10-024-632-3 Sequence 3, App
US-10-325-066A-591 Sequence 591, App
US-10-374-780A-339 Sequence 339, App
US-10-059-911-4 Sequence 4, App
US-10-225-066A-381 Sequence 381, App
US-10-374-780A-345 Sequence 345, App
US-10-183-687-337 Sequence 337, App

ALIGNMENTS

RESULT 1

US-10-183-687-355
Sequence 355, Application US/10183687
Publication No. US20030204870A1
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Allen, William B.
APPLICANT: Cahoon, Rebecca
APPLICANT: Epelbaum, Sabine
APPLICANT: Famodu, Omolayo O.
APPLICANT: Harvell, Leslie T.
APPLICANT: Jones, Todd
APPLICANT: Kinney, Tony
APPLICANT: Klein, Ted
APPLICANT: Li, Changjiang
APPLICANT: Oliveira, Igor Cunha
APPLICANT: Sakai, Hajime
APPLICANT: Shet, Bo
APPLICANT: Tarczynski, Mitchell C.
TITLE OF INVENTION: Alteration Of Oil Traits In Plants
FILE REFERENCE: B4458 US NA
CURRENT APPLICATION NUMBER: US/10/183,687
CURRENT FILING DATE: 2002-06-27
PRIOR APPLICATION NUMBER: 60/301,913
PRIOR FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 532
SOFTWARE: Microsoft Office 97
SEQ ID NO 355
LENGTH: 2463
TYPE: DNA
ORGANISM: Glycine max
US-10-183-687-355

Query Match 20.9%; Score 420.6; DB 13; Length 2463;
Best Local Similarity 76.4%; Pred. No. 1.8e-105;
Matches 516; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

; APPLICANT: Oliveira, Igor Cunha
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Shen, Bo
 ; APPLICANT: Tarczynski, Mitchell C.
 ; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
 ; FILE REFERENCE: BSI458 US NA
 ; CURRENT APPLICATION NUMBER: US/10/183,687
 ; CURRENT FILING DATE: 2002-06-27
 ; PRIOR APPLICATION NUMBER: 60/301,913
 ; PRIOR FILING DATE: 2001-06-29
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 351
 ; LENGTH: 2598
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-183-687-351

Query Match 20.5%; Score 413.6; DB 13; Length 2598;
 Best Local Similarity 79.7%; Pred. No. 1.6e-103; Indels 0; Gaps 0;
 Matches 488; Conservative 0; Mismatches 124;

QY	695	CGTTGAAGCTACCGGAGAAACTATTGAGAGTTTTGGACAGAGACCTCTATATACG	754
DB	773	CATTGAAGCTACCGGAGAACTTATGACCTTTTGGACAGAGAACTTCTATCTACG	832
QY	755	CGGTGTTACAGGCTCGGTGGACAGAGATATGAGGCACATTTATGGATAATAGTGG	814
DB	833	TGGTGTAAAGGCTATAGTGGAGCGGAGGTACGAGCTCCTCTGGGTAATAGTGG	892
QY	815	TAAAGAGAGGCGAAACCGCGAAAGAGACAAAGTTTATTTGGGAGGTTATGACAAAGA	874
DB	893	TAGAAGAGAGGACAGACTCGCAAGAGGAGCAAGTTTACTTGGGTGTTATGATAAGA	952
QY	875	AGAAAGAGCAGCTAGGCTTATGATTTAGCGCACTCAAGTATTTGGGAAACCACTAC	934
DB	953	AGAAAGAGCAGCTAGACCTACGATTTGGCAGCACTAAATCTGGGAAACCACTAC	1012
QY	935	TACTAACTTCCCATGAGCGAATATGAAAGAGGTAGAGAGATGAAGCAGACATGAC	994
DB	1013	AACAAATTTTCCAATTAGCCACTATGAGAAAGAGTTGGAAGAAATGAAGCAGATG	1072
QY	995	GCAGAGTATGTCCTCACTCGCGAGGAAAGTATGTTCTCTCGTGGTGCATCGAT	1054
DB	1073	GCAAGAGTATGTCGTCTTGAAGAGAGTATGTTCTCTCGTGGTGCATCGAT	1132
QY	1055	TTATCGTGGAGTAACAAGACATCAACCAATGGAAGATGCGCAAGCTTAGGATAGG	1114
DB	1133	TTATCGAGGAGTACGAGACACACCAATGGAAGTGGCAAGCGAGGATTTGAAGCT	1192
QY	1115	CGCGGTAAACAAGACCTCTACTTGGAACTTTTGGACACACAGAGAGAGCTGCAG	1174
DB	1193	TGCTGGCAACAAGATCTTACTTGGAACTTTTGGACACACAGAGAGAGCTGCAG	1252
QY	1175	ATACGACATTTGGGCGCATCAAAATTCAGAGATTTAACCGCAGTGACTTAACCT	1234
DB	1253	ATATGATGAGAGCAATCAAAATTCGAGGACTTAAGTGTGTTTACAACTTTGAC	1312
QY	1235	CAGTACAGCTTAAAGCACTCTCGAAGCCCTAGTCTTCTTATTTGGTAGCGCGCAA	1294
DB	1313	CAGATATGAGCTGAAAGCATACTTTGAGAGCACCCTTTGCCAATAGTGGTGTG	1372
QY	1295	ACGTCTCAAGGA	1306
DB	1373	CGGTTTGAAGGA	1384

RESULT 4

US-10-183-687-361
 ; Sequence 361, Application US/10183687
 ; Publication No. US20030204870A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve

; APPLICANT: Allen, William B.
 ; APPLICANT: Caboon, Rebecca
 ; APPLICANT: Ebelbaum, Sabine
 ; APPLICANT: Fomodu, Omolayo O.
 ; APPLICANT: Harvell, Leslie T.
 ; APPLICANT: Jones, Todd
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Klein, Ted
 ; APPLICANT: Li, Changjiang
 ; APPLICANT: Oliveira, Igor Cunha
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Shen, Bo
 ; APPLICANT: Tarczynski, Mitchell C.
 ; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
 ; FILE REFERENCE: BSI458 US NA
 ; CURRENT APPLICATION NUMBER: US/10/183,687
 ; CURRENT FILING DATE: 2002-06-27
 ; PRIOR APPLICATION NUMBER: 60/301,913
 ; PRIOR FILING DATE: 2001-06-29
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 361
 ; LENGTH: 1990
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-183-687-361

Query Match 19.2%; Score 385.8; DB 13; Length 1990;
 Best Local Similarity 76.9%; Pred. No. 7.2e-96; Indels 3; Gaps 1;
 Matches 484; Conservative 0; Mismatches 142;

QY	692	TAGCGTTGAAGCTACACCGGAGAAACTATTGAGAGTTTTGGACAGAGACCTCTATATA	751
DB	560	TACTGTTGAAGTTGCACCTAGAGAACTTTGGATACATTCGGGACAGAACTCCATATA	619
QY	752	CGCGGTGTTTACAGGCTCGGTGGACAGAGATATGAGGCACATTTATGGATATAG	811
DB	620	TCGTGGAGTAACTCGACATAGATGGAGGTATGAAGCTCATCTTTGGGATATAG	679
QY	812	TTGTAAAGAGAGGCGCAACCGCGCAAGAGAAACAAAGTTTATTTGGGAGGTTATGACAA	871
DB	680	CTGTAGAAGGAGGAGCCCAATCAAGAAAGAGCAAGTTTATTTGGTGGATATGATAA	739
QY	872	AGAAAGAAAGCAGCTAGGCTTATGATTTAGCCGCACTCAAGTATTTGGGAAACCAAC	931
DB	740	AGAAAGAAAGCAGCTAGGCTTATGATTTAGCTGCACTGAAGTACTGGGGGACATCCAC	799
QY	932	TACTACTTAACTTCCCATGAGCGAATATGAAAGAGGTAGAGAGATGAAGCAGATGAC	991
DB	800	CACTACCAACTTTTCCCAATTAGCAACTATGAGAAGGAATTTGGATGAATGAACACATGAC	859
QY	992	AAGCAAGATATTTGGCTCACTGCGCAGGAAAGTAGTGGTTTCTCTCGTGGTGCATC	1051
DB	860	GAGACAGAAATTTGTTGCGCCATTAGAAAGGAAAGCACTGGTTTCTCCAGGGGTGCATC	919
QY	1052	GATTTATCGTGGAGTAACAAGACATCAACCAATGGAAGATGGAAGCTTAGGATAGGAAG	1111
DB	920	AATGATATCGTGGAGTTTACAGGCACTCAACCAACGAGATGGAAGCAAGGATTGGCAG	979
QY	1112	AGTCGCGGTAAACAAGACCTCTACTTGGGAACCTTTTGGACACAGAGAGAGCTCCAGA	1171
DB	980	AGTTGCAAGAAACAAAGATCTTTTACTTGGAACTTTTCTAGTACTGAGGAAGAGCTGCAGA	1039
QY	1172	GGCATACGACATTTGGGCGCATCAAAITTCAGAGATTAACCGCAGTCACTAACTTCACAT	1231
DB	1040	AGCATACGACATAGCAGCGATTAAGTTTCAAGGTCTCAACGCTGTGCACAACTTTGACAT	1099
QY	1232	GAACAGATCAACGTTTAAAGCAATCTCGAAGCCCTAGTCTTCTCTAT---TGGTAGCCG	1288
DB	1100	GAGCCGCTACGAGTGAAGGCCAATTTTGAAGCAACACTCTCCCAATAGGAGGAGCCG	1159
QY	1289	CGCAAAACGCTCTCAAGGAGGCTTAACCGTC	1317

Db 1160 TCGAAGCGCTCTGAAAGAGCTCAAGCTC 1188

RESULT 5
US-10-374-780A-1738
; Sequence 1738, Application US/10374780A
; Publication No. US2004001927A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Cressman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James
; APPLICANT: Brown, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1738
; LENGTH: 2088
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G1793
US-10-374-780A-1738

Query Match 18.9%; Score 381; DB 16; Length 2088;
Best Local Similarity 76.5%; Pred. No. 1.6e-94;
Matches 481; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

QY 692 TAGCGTTGAAGCTACACCGAAGAAACTATTGAGAGCTTTTGACAGAGGACGCTCTATATA 751
Db 567 TACTGTTGAAGTTGCACCTAGAAGAACTTTGGATACATTCGGCGACAGAACTCATATA 626
QY 752 CCGCGGTGTTACAGGCTCGGTGGACAGAGATATGAGGCATTTATGGATATAG 811
Db 627 TCGTGGAGTAACTCGACATAGATGGACTGGAAGGTATGAAGCTCATCTTTGGGATAATAG 686
QY 812 TTGTAAGAGAGAGGCGCAACCGCGCAAGGAGACAAAGTTTATTTGGGAGGTTATGACAA 871
Db 687 CTGTAGAAGGAGGAGGCAATCAAGAAAGGACGCCCAAGTTTATTTGGGTGATATGATA 746
QY 872 AGAAGAAAGAGCTAGGCTTATGATTTAGCCGACCTCAAGTATTGGGGAACCCAC 931

Db 747 AGAAGAGAAAGCAGCTAGGGCTTATGATTTAGCTGCACTGAAGTACTGGGGACATCCAC 806
QY 932 TACTACTAACTTCCCATGAGCGAATATGAAAAGAGGTAGAGAGATGAAGCATGAC 991
Db 807 CACTACCAACTTCCNAATTAGTAATCTATGAGAAGAAATGGATGAATGAACACATGAC 866
QY 992 AAGGCAAGAGTATGTTGCCTCACTCGCGAGGAAAAAGTAGTGTCTCTCGTGTGTCATC 1051
Db 867 GCGACAAGAAATTTGTTGCTGCCATTAGAAAGGAAAAGCAGTGTCTTCTCCAGGGGTGCATC 926
QY 1052 GATTTATCGTGGAGTAAAGACATCAACACATCGAAGATGGCAAGCTAGATAGGAAG 1111
Db 927 AATGTATCGTGGAGTTACAAGGCATCAACACACGGAAGATGGCAAGAAATGGCAG 986
QY 1112 AGTCGCGGTAAACAAAGACCTCTACTTGGGAACCTTTTGGCACACAAAGAAAGCTGCAG 1171
Db 987 AGTTGCGAGGAAACAAAGATCTTTACTTGGGAACCTTTCAGTACTGAGAAGAGGCTGCTGA 1046
QY 1172 GGCATACGACATTTGGGCGCATCAAAATTCAGAGGATTAACCGAGTGAATACTTCGACAT 1231
Db 1047 AGCATACGACATAGCTGCGATAAAGTTTCAGAGGTCTCAACGCTGTCAAAAACCTTTGACAT 1106
QY 1232 GAACAGATACAACTGTTAAAGCAATCTCTGAAAGCCCTAGTCTTCTAT---TGGTAGCGC 1288
Db 1107 GAGCGCTACGACGTGAAGGCCATCTTGAAGCAACACTCTCCCATAGGAGGAGGC 1166
QY 1289 CGCAAAAGCTCTCAAGGAGGCTAACCGTC 1317
Db 1167 TCGAAAGCGTCTGAAAGAAAGCTCAAGCTC 1195

RESULT 6
US-10-424-599-91337
; Sequence 91337, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 91337
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53468C.1
US-10-424-599-91337

Query Match 18.9%; Score 380.8; DB 13; Length 2169;
Best Local Similarity 75.8%; Pred. No. 1.9e-94;
Matches 485; Conservative 0; Mismatches 152; Indels 3; Gaps 1;

QY 681 ACTATTGATGATACGCTTGAAGCTACACCGAAGAAACTATTGAGAGTCTTTGGACAGAGG 740
Db 568 ACTACTACTACTCTGTTGAAGTTGACCTTAGAAGACATTTGGATACATTCGGGCGAGAG 627
QY 741 ACCTCTATATACCGCGGTGTTCAAGGCATCGGTGGACAGAAAGATATAGGCACATTTA 800
Db 628 ACATCCATATATCTGCGAGTAAACCGCACATAGATGGACTGGAAGGTATGAAGCTCATCT 687
QY 801 TGGGATATAGTTGTAAGAGAGAGGCGCAACCGCGCAAGGAGAGACAACTTTATTTGGGA 860
Db 688 TGGGATATAGTGTAGAGGAGGCGCAATCAAGAAAGAGCGCCCAAGTTTATTTGGGT 747
QY 861 GGTATCAACAAAGAAAAAGCAGCTAGGCTTATGATTTAGCCGACCTCAAGTATTGG 920
Db 748 GGATATGATAAGAAAAAGCAGCTAGGCTTATGATTTAGCTGCACTGAAGTACTGG 807

QY 921 GGAACCCACTACTACTACTTCCCATGAGCAATATGAAAGAGGTAGAGATG 980
DB 808 GGGACATCCACCTACTCCACTTCCAAATAGTAATATGAGAGAAATGGATGAATG 867
QY 981 AAGCACATGACAAAGGAGATGATGTTGCTCCTACTCGCAGGAAAAGTGTCTCT 1040
DB 868 AAACACATGACGCGACAAAGAAATGTTGCTGCCATTAGMAGGAAAAGAGTGTCTCC 927
QY 1041 CGTGTGCATCGATTTATCGTGGAGTAAACAGACATCACCACATGGAAGATGGCAAGCT 1100
DB 928 AGGGTGCATCAATGTTATCGTGGAGTTACAGGCAATCACCACACGGAAGATGGCAAGCA 987
QY 1101 AGATAGAGAGATGCGCGGTAAACAAAGACCTTACTTTGGAACTTTTGGCACACAGAA 1160
DB 988 AGAATGCGAGATGTCAGGAAAACAAAGATCTTTACTTTGGAACTTTTCACTACTGAAGAA 1047
QY 1161 GAAGTCGAGAGGATGAGCATGCGGCATCAAAATTCAGAGGATTAACCGCAGTACT 1220
DB 1048 GAGGTGCTGAAGCATGACATAGCTGCGATTAAGTTCAAGGTCTCAACGCTGTCACA 1107
QY 1221 AACTTCGACATGAACAGATCAACGTTTAAAGCAATCTCGAAAGCCCTAGTCTTCCAT- 1279
DB 1108 AACTTTGACATGAGCGCTAGCAGCTGAAAGCCATCTTTGAAAGCAACACTCTCCCAATA 1167
QY 1280 --TGTAGCGCGCAAAAGCTCTCAAGAGGCTTAACCGTC 1317
DB 1168 GGAGAGAGAGTGCAAAAGCGTCTGAAGAAGCTCAAGCTC 1207

RESULT 7

US-10-225-066A-365
; Sequence 365, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MEI0036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 365
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-365

Query Match 18.7%; Score 377.4; DB 13; Length 1818;
Best Local Similarity 76.7%; Pred. No. 1.5e-93;
Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

QY 702 GCTACACCGAAGAACTATTGAGAGTTTGGACAGAGACCTCTATATACCGCGTGT 761
DB 554 GCCACGCCAAGAGCGTGCATTGGACATTTTCGCAACAGAACCTCGATCTATCGTGTGTC 613
QY 762 ACAAGCATCGTGGAGAGAGATATGAGGCACATTTATGGGATAATAGTTTGTAAAGA 821
DB 614 ACAAGCATCGATGAGTGTGTCATATGAGGCTCATCTATGGGATAATAGTTTGTAAAG 673
QY 822 GAAGCCCAACCGCGCAAGAGACACAAAGTTTATTTGGAGGTTATGACAAAGAGAAAAA 881
DB 674 GAAGCCCGTCTAGGAAGAGACAAAGTTTATTTGGTGGATATGACAAAGAGATAAA 733
QY 882 GCAGCTAGGCTTTATGATTTAGCGCCTCAAGTATTTGGGAAACACCACTACTACTAAC 941
DB 734 GCAGCAAGATCATATGATCTAGCTGCACCTTAGTCTGGGCTCTTCAACTACTACTAAT 793
QY 942 TTCCCCATGAGCAATATGAAAGAGGTAGAGATCAAGCACAATGACAAAGGCAAGAG 1001
DB 794 TTCCCCATTTCAAACTACGAGAAAGAGTAGAGAAATGAGCAATGACGACGACAGAG 853
QY 1002 TATGTTCCCTCACTGCGCAGGAAAAGTAGTGGTTTCTCTGTTGGTGCATCGATTTATCGT 1061
DB 854 TTGTTGGCTGCCATTAGAAGGAAAAGTAGTGGATTTTCGAGAGGCGCTTCGATGTATCGA 913
QY 1062 GGATTAACAGACATCACCACCAATGGAAGATGCAAGATGCAAGATGCAAGATGCAAGATGCAAGAT 1121
DB 914 GGATTAACAGGATCACCACCAATGGAAGATGCAAGATGCAAGATGCAAGATGCAAGATGCAAGAT 973
QY 1122 AACAAAGACCTTACTTTGGAACTTTTGGCACAACAAGAAAGCTCGCAGGAGGATACGAC 1181
DB 974 AACAAAGACCTTACTTTGGAACTTTTGGCACAACAAGAAAGCTCGCAGGAGGATACGAT 1033
QY 1182 ATTGCGGCCATCAAAATTCAGAGGATTAACCGCAGTCACTAACTTCGACATGACAGATAC 1241
DB 1034 ATAGTCAATTAAGTTTAGAGGACTTAATGCAAGTACCAACTTCGAGATCAACCGGTAC 1093
QY 1242 AACCTTAAAGCAATCTCGAAAGCCCTAGTCTTCCAT--TGGTAGCGCCGCAAAAGCT 1298
DB 1094 GACGTGAAAGCCATTCTAGAGAGTAGCACTCTTCCATCGGAGAGGCGCAGCTAAACGG 1153
QY 1299 CTCGAAGGAGGCTAACCGTC 1317
DB 1154 CTCGAAGGAGGCTAACCGTC 1172

RESULT 8

US-10-225-066A-855
; Sequence 855, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MEI0036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-12-05
 ; PRIOR APPLICATION NUMBER: 60/338,692
 ; PRIOR FILING DATE: 2001-12-11
 ; PRIOR APPLICATION NUMBER: 10/171,468
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 1122
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 855
 ; LENGTH: 1818
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-225-066A-855

Query Match 18.7%; Score 377.4; DB 13; Length 1818;
 Best Local Similarity 76.7%; Pred. No. 1.5e-93;
 Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

QY	702	GCTACACCGAAGAACTATGAGAGTTTGGACAGAGGCTCTATATACCGGGTGT 761
DB	554	GCACGCCAAGACGTGCTATGGACACTTCGACAAACCTCGATCTATCGTGGTGC 613
QY	762	ACAAGGCATCGTGGACAGGAGATATGAGGCACATTTATGGGATAATAGTTTAAAGA 821
DB	614	ACAAGACATCGATGACTGGTGCATATGAGGCTCATCTATGGGATAATAGTTTGAAGG 673
QY	822	GAAGCCAAACCGCAAGAGACAGTTTATTTGGAGGTTATGACAAAGAGAAAA 881
DB	674	GAAGGCCATCTAGAAAGAGAGACAGTTTATTTGGGTTGATATGACAAAGAGATAA 733
QY	882	GCAGCTAGGGCTTATGATTTAGCCGCACTCAAGTATTTGGGAAACCACTACTTAAAC 941
DB	734	GCAGCAAGATCATATGATCTAGCTGCACTTAAGTACTGGGCTCTTCAACTACTTAAAT 793
QY	942	TTCCCATGAGCGAATATGAAAAGAGGTAGAGACATGAGCAGATGACAGGCAAG 1001
DB	794	TTCCCATTTACAACTACGAGAAAGTAGAGAAATGAAAGCAGATGACAGACAGAG 853
QY	1002	TATGTTGGCTCTACTCGCAGGAAAGTAGTGTCTCTCTGTTGTCATGATTTATCGT 1061
DB	854	TTCTGGTGGCTGCAATAGAAAGAAAGTAGTGTGATTTTCGAGAGGCGCTTCGATGTA 913
QY	1062	GGAGTAAAGCAATCCTCGAAGCCCTAGTCTTCTCTAT---TGGTAGCGCGCAAAACGT 1121
DB	914	GGAGTTACAAAGTTTAGAGACTTAATGAGTGAACCACTTCGAGATCAACCGGTAC 973
QY	1122	AACAAAGACCTCTACTTGGGAACCTTTGGCACACAAAGAAAGCTGCAGAGGCATACAC 1181
DB	974	AACAAAGACCTCTACTTGGGAACCTTTTAGCACTGAGGAAGAGCAGCAAGCTTACGAT 1033
QY	1182	ATTGGGCGCATCAATTCAGAGGATTAACCGCAGTACTACTTTCGACATGACAGATAC 1241
DB	1034	ATAGCTGCATTAAGTTTAGAGACTTAATGAGTGAACCACTTCGAGATCAACCGGTAC 1093
QY	1242	AACGTTAAAGCAATCCTCGAAGCCCTAGTCTTCTCTAT---TGGTAGCGCGCAAAACGT 1298
DB	1094	GACGTGAAAGCCATCTAGAGAGTAGCACTCTTCCCATCGGAGGAGCGGAGCTAAACGG 1153
QY	1299	CTCAAGGAGGCTAAACGGTC 1317
DB	1154	CTCAAGAGAGCTCAAGCTC 1172

RESULT 9
 US-10-225-066A-911
 ; Sequence 911, Application US/10225066A
 ; Publication No. US20030226173A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendel Biotechnology, Inc.
 ; APPLICANT: RATCLIFFE, Oliver
 ; APPLICANT: RIECHMANN, Jose Luis
 ; APPLICANT: ADAM, Luc J
 ; APPLICANT: DUBELL, Arnold T
 ; APPLICANT: HEARD, Jacqueline E

; APPLICANT: PILGRIM, Marsha L
 ; APPLICANT: JIANG, Cai-Zhong
 ; APPLICANT: REUBER, T. Lynne
 ; APPLICANT: CREELMAN, Robert A
 ; APPLICANT: PINEDA, Omaira
 ; APPLICANT: YU, Guo-Liang
 ; APPLICANT: BROWN, Pierre E
 ; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
 ; FILE REFERENCE: MB10036-2 US
 ; CURRENT APPLICATION NUMBER: US/10/225,066A
 ; CURRENT FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: 09/837,444
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: 60/310,847
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: 60/336,049
 ; PRIOR FILING DATE: 2001-12-05
 ; PRIOR APPLICATION NUMBER: 60/338,692
 ; PRIOR FILING DATE: 2001-12-11
 ; PRIOR APPLICATION NUMBER: 10/171,468
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 1122
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 911
 ; LENGTH: 1818
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-10-225-066A-911

Query Match 18.7%; Score 377.4; DB 13; Length 1818;
 Best Local Similarity 76.7%; Pred. No. 1.5e-93;
 Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

QY	702	GCTACACCGAAGAACTATGAGAGTTTGGACAGAGGCTCTATATACCGGGTGT 761
DB	554	GCACGCCAAGACGTGCTATGGACACTTCGACAAACCTCGATCTATCGTGGTGC 613
QY	762	ACAAGGCATCGTGGACAGGAGATATGAGGCACATTTATGGGATAATAGTTTAAAGA 821
DB	614	ACAAGACATCGATGACTGGTGCATATGAGGCTCATCTATGGGATAATAGTTTGAAGG 673
QY	822	GAAGCCAAACCGCAAGAGACAGTTTATTTGGAGGTTATGACAAAGAGAAAA 881
DB	674	GAAGGCCATCTAGAAAGAGAGACAGTTTATTTGGGTTGATATGACAAAGAGATAA 733
QY	882	GCAGCTAGGGCTTATGATTTAGCCGCACTCAAGTATTTGGGAAACCACTACTTAAAC 941
DB	734	GCAGCAAGATCATATGATCTAGCTGCACTTAAGTACTGGGCTCTTCAACTACTTAAAT 793
QY	942	TTCCCATGAGCGAATATGAAAAGAGGTAGAGACATGAGCAGATGACAGGCAAG 1001
DB	794	TTCCCATTTACAACTACGAGAAAGTAGAGAAATGAAAGCAGATGACAGACAGAG 853
QY	1002	TATGTTGGCTCTACTCGCAGGAAAGTAGTGTCTCTCTGTTGTCATGATTTATCGT 1061
DB	854	TTCTGGTGGCTGCAATAGAAAGAAAGTAGTGTGATTTTCGAGAGGCGCTTCGATGTA 913
QY	1062	GGAGTAAAGCAATCCTCGAAGCCCTAGTCTTCTCTAT---TGGTAGCGCGCAAAACGT 1121
DB	914	GGAGTTACAAAGTTTAGAGACTTAATGAGTGAACCACTTCGAGATCAACCGGTAC 973
QY	1122	AACAAAGACCTCTACTTGGGAACCTTTGGCACACAAAGAAAGCTGCAGAGGCATACAC 1181
DB	974	AACAAAGACCTCTACTTGGGAACCTTTTAGCACTGAGGAAGAGCAGCAAGCTTACGAT 1033
QY	1182	ATTGGGCGCATCAATTCAGAGGATTAACCGCAGTACTACTTTCGACATGACAGATAC 1241
DB	1034	ATAGCTGCATTAAGTTTAGAGACTTAATGAGTGAACCACTTCGAGATCAACCGGTAC 1093
QY	1242	AACGTTAAAGCAATCCTCGAAGCCCTAGTCTTCTCTAT---TGGTAGCGCGCAAAACGT 1298
DB	1094	GACGTGAAAGCCATCTAGAGAGTAGCACTCTTCCCATCGGAGGAGCGGAGCTAAACGG 1153

QY 1299 CTCAAGGAGGCTAACCGTC 1317
|||
DB 1154 CTCAAGAGAGCTCAAGCTC 1172

RESULT 10

US-10-374-780A-333

; Sequence 333, Application US/10374780A

; Publication No. US2004001927A1

; GENERAL INFORMATION:

; APPLICANT: Sherman, Bradley K

; APPLICANT: Riechmann, Jose Luis

; APPLICANT: Jiang, Cai-Zhong

; APPLICANT: Heard, Jacqueline E

; APPLICANT: Haake, Volker

; APPLICANT: Creelman, Robert A

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Adam, Luc J

; APPLICANT: Reuber, T. Lynne

; APPLICANT: Keddle, James

; APPLICANT: Brown, Pierre E

; APPLICANT: Pilgrim, Marsha L

; APPLICANT: Dubell III, Arnold T

; APPLICANT: Pineda, Omaira

; APPLICANT: Yu, Guo-Liang

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

; FILE REFERENCE: MBI-0047 CIP

; CURRENT APPLICATION NUMBER: US/10/374,780A

; CURRENT FILING DATE: 2003-02-25

; PRIOR APPLICATION NUMBER: 09/837,944

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/310,847

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 09/934,455

; PRIOR FILING DATE: 2001-08-22

; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-11-19

; PRIOR APPLICATION NUMBER: 60/338,692

; PRIOR FILING DATE: 2001-12-11

; PRIOR APPLICATION NUMBER: 10/171,468

; PRIOR FILING DATE: 2002-06-14

; PRIOR APPLICATION NUMBER: 10/225,066

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 10/225,067

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 10/225,068

; PRIOR FILING DATE: 2002-08-09

; NUMBER OF SEQ ID NOS: 2906

; SOFTWARE: Patent in version 3.2

; SEQ ID NO 333

; LENGTH: 1818

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: G1793

US-10-374-780A-333

Query Match 18.7%; Score 377.4; DB 16; Length 1818;

Best Local Similarity 76.7%; Pred. No. 1.5e-93;

Matches 475; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

QY 702 GCTACACGGAAGAAAACCTATTGAGAGTTTGGACAGAGGAGCTCTATATACCGCGTGT 761

DB 554 GCCACGCCAAGACGTGATGGACACTTCCGACACGACCTCGATCTATCGTGTGTC 613

QY 762 ACAAGGATCCGTCGACAGGAAGATATGAGGCACATTATGGGATATATAGTTGTAAGA 821

DB 614 ACAAGATCATGAGCTGGTCGATATGAGGCTCACTATGGGATATAGTTGTAGAAG 673

QY 822 GAAGCCCAAGCGCCAAAGGAAGCAAGATTATTTGGAGGTTATGACAAAGAGAAAAA 881

DB 674 GAAGCCGAGTCTAGGAAGGAAGCAAGATTATTTGGTGGTATGACAAAGAGATAAA 733

QY 882 GCAGCTAGGGCTTTATGATTTAGCCGCACTCAAGTATTGGGGAACCACTACTACTAAC 941
|||
DB 734 GCAGAGATCATATGATCTAGCTGCACCTAAGTACTGGGTCCTTCAACTACTACTAT 793
|||
QY 942 TTCCCATGAGCGAATATGAAGAAGAGGTAGAGAAGATGAAGCATGACAGGCGAAG 1001
|||
DB 794 TTCCCATTTACAAACACGAGAAAGAGTAGAGAAATGAAGCATGACAGAGCAAGAG 853
|||
QY 1002 TATGTTGCCTCACTCGCGCAGGAAAAGTAGTGGTTTTCTCTCGTGGTGCATCGATTTATCGT 1061
|||
DB 854 TTGTTGCTGCCATTTAGAAGGAAAAGTAGTGGATTTCGAGAGGCGCTTCGATGATCGA 913
|||
QY 1062 GGAGTAACAGACATCACCAACATGGAAGATGGCAGCTAGGATAGGAGAGCTCCCGGT 1121
|||
DB 914 GGAGTTACAAGGCATCACCAACATGGAAGATGGCAGGAGGATCGCGGATCGCGCGA 973
|||
QY 1122 AACAAAGACCTCTACTTTGGGAACCTTTTGGCACACAAGAAGAAGCTGCAGAGGCATACGAC 1181
|||
DB 974 AACAAAGACCTCTACTTTGGGAACCTTTTGGCACACAAGAAGAAGCAGCAGAAGCTTACGAT 1033
|||
QY 1182 ATTGCGGCCATCAATTCAGAGGATTAACCGCAGTGACTTAACCTCGACATGAACAGATAC 1241
|||
DB 1034 ATAGCTGCAATAAAGTTTAGAGGACTTAATGCAAGTGACCACTTCGAGATCAACCGGTAC 1093
|||
QY 1242 AACGTTAAGCAATCCCTCGAAAGCCCTAGTCTTCTCTAT---TGGTAGCCCGCAAAACGT 1298
|||
DB 1094 GAGTGAAGCCCATTTAGAGAGTAGCAGCTCTTCCCATCGGAGGAGGCGCAGCTAAACGG 1153
|||
QY 1299 CTCAGAGGAGCTAACCGTC 1317
|||
DB 1154 CTCAAGAGAGCTCAAGCTC 1172
|||

RESULT 11

US-10-425-114-20762

; Sequence 20762, Application US/10425114

; Publication No. US20040034888A1

; GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong

; APPLICANT: Zhou, Yihua

; APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steven E

; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 20762

; LENGTH: 794

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3245-185-F4_FLI

US-10-425-114-20762

Query Match

Best Local Similarity 18.1%; Score 364.4; DB 13; Length 794;

Matches 449; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 720 ATTGAGAGTTTGGACAGAGGAGCTCTATATACCGCGGTGTACAGGATCGGTGGACA 779

DB 1 ATCGACACGTTCCGGCAAGGACCTCTATATATCGAGGTGTAAACAGGATAGATGGACA 60

QY 780 GGAAGATATCAGGCACATTTATGGGATAATAGTTGTAAAGAGAAGGCCAAACGCGCAAA 839

DB 61 GGGCGGTATGAGGCTCATCTATGCGATATAGTTGTAGAGGAGGAGGAGGATGCGAAG 120

QY 840 GGAAGCAAGTTTATTTGGAGGTTATGACAAAGAGAAAAAGCAGCTAGGCGTTATGAT 899

DB 121 GGTAGGCAAGTTTACCTTGGTGGTCTATGACAGGAGGACAAAGCAGCAGAGGCTTACGAT 180

APPLICANT: Cac, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: B1458 US NA
CURRENT APPLICATION NUMBER: US/10/183,687
CURRENT FILING DATE: 2002-06-27
PRIOR APPLICATION NUMBER: 60/301,913
PRIOR FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 532
SOFTWARE: Microsoft Office 97
SEQ ID NO 486
LENGTH: 1935
TYPE: DNA
ORGANISM: Gossypium hirsutum
US-10-183-687-486

Query Match 17.7%; Score 357.4; DB 13; Length 1935;
Best Local Similarity 73.8%; Pred. No. 5.e-88;
Matches 454; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 658 ACAACGTTTGGCCCAAGGAAGACTATTGATGATAGCGTTGAAGCTACACGGAAGAAA 717
Db 442 AATCGATATTACAAACACACGACACAGCAAGCACTATAGCTGAGCGCTGCACCCAAAGAT 501
QY 718 CTATTGAGAGTTTGGACAGAGAGCTCTATATACCGCGTGTTCACAGGCTCGTGGGA 777
Db 502 CTGTTGATATCTTTGGCCCAACGTACCTCGATCTACCGGGGTGTCCAGGCGATAGTGA 561
QY 778 CAGGAAGATATGAGGCACATTTATGGGTAATAGTTTAAAGAGAGAGCCCAACGCCCA 837
Db 562 CTGGAAGGTATGAGCTCTCTGTTGGCAATAGTTTCAGAGAGAGAGCCCAAGTAGGA 621
QY 838 AAGGAAGCAAGTTTATTTGGGAGTTTATGACAAAGAGAGAGAGAGAGAGAGAGAGAG 897
Db 622 AAGGAAGCAAGTTTATTTGGGTGGTGTATGACAAAGGAAGATAAGCTGCAAGAGCTTATG 681
QY 898 ATTTAGCCCACTCAAGTATTTGGGAAACCACTACTACTACTACTACTACTACTACTACT 957
Db 682 ATCTTGGCGCTCTCAAGTATCTGGGTTCGACCACTACTACTACTACTACTACTACTACT 741
QY 958 ATGAAAAGAGGTAGAGAGATGAAGCAATGACAAAGCAAGAGTATGTTGCTCTACTGC 1017
Db 742 ATGAGAAGAGCTCGAAGAGATGAAGCAATGACAAAGCAAGTTCGTTGCTTCTCTCC 801
QY 1018 GCAGGAAGAGTATGTTTCTCTCGTGGTGCATGATTTATCGTGAGTAAGCAAGACATC 1077
Db 802 GGAAGGAAGTATGGAATTTTCAGGGGTGCTTCAATTTACAGAGAGTGAAGAGCATC 861
QY 1078 ACCAAGATGAAGATGCAAGCTAGGATAGGAAGAGTTCGCGGTAAACAAAGACCTTACT 1137
Db 862 ATCAACATGTAGGTGCAAGCAAGATTTGGAAGAGTTGCAAGCAACAAAGATCTCTATC 921
QY 1138 TGGGAATTTTGGACACAAAGAGAGCTGCAGAGCATACGACATTCGGCCCATCAAT 1197
Db 922 TTGSCACATTTAGCACCAAGAGAGAGCAGCTGAAGCCTATGATATTGCAAGCAATCAAGT 981
QY 1198 TCAGAGGATTAACCGCAGTCACTAACTTCGACATGAACAGATCAACAGTTTAAAGCAATCC 1257
Db 982 TTAGAGCTTAAACCGGTTACAAATTTGATATGAGCCCTACGATGAACAGCATTTG 1041
QY 1258 TCGAAAGCCCTAGTC 1272
Db 1042 CAAACAGCAATCTTC 1056

RESULT 14
US-10-425-114-14859
Sequence 14859, Application US/10425114
Publication No. US20040034889A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kowalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.

APPLICANT: Cac, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: B1458 US NA
CURRENT APPLICATION NUMBER: US/10/183,687
CURRENT FILING DATE: 2002-06-27
PRIOR APPLICATION NUMBER: 60/301,913
PRIOR FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 532
SOFTWARE: Microsoft Office 97
SEQ ID NO 486
LENGTH: 1935
TYPE: DNA
ORGANISM: Gossypium hirsutum
US-10-183-687-486

Query Match 17.7%; Score 357; DB 13; Length 1479;
Best Local Similarity 76.4%; Pred. No. 6.e-88;
Matches 438; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 700 AAGTACACCGAAGAAAACCTATTGAGAGTTTGGACAGAGAGCTCTATATACCGCGTG 759
Db 38 AACCTCTCCAAAGAAAACCGTCGACACCTTCGGCCAAACGACCTCCATCTACCGCGCG 97
QY 760 TTACAAGGCTATCGGTGGACAGAGATATGAGGCACATTTATGGGTAATAGTTGTAAAA 819
Db 98 TCACCCACATAGATGACGGAAGATACGAAGCTCATCTATGGCAATAGTTGTAGAA 157
QY 820 GAGAAGGCCAAACCGCAAGAGAAAGCAAGTTTATTTGGGAGGTATGACAAAGAGAAA 879
Db 158 GAGAAGGCCAAAGCAGGAAAGCAAGTTTACCTGGGTGGTTATGACAAAGAGAGATA 217
QY 880 AAGCAGCTAGGCTTATGATTTAGCCGACCTCAAGTATTTGGGAAACCACTACTACTA 939
Db 218 AGGAGAGCCAGGCTTACGATCTCGAGCTCTCAAGTACTGGGTCCAACCTACCACCA 277
QY 940 ACTTCCCATGAGCAATATGAAAAGAGGTAGAAAGATGAAGCAATGACAAAGCAAG 999
Db 278 ACTTCCCATTTCCAACTATGAGAAGGAACCTGGAAGAGATGAAGCAATGACAAAGCA 337
QY 1000 AGTATGTTGCTCTACTCGCAGGAAAGTATGTTTCTCTCGTGGTGCATCGATTTATC 1059
Db 338 AGTTGTTGCTTCTACGAAGAGAGCAGTGGTTTCTTAGGGGGCTCTATATACA 397
QY 1060 GTGAGTAACAAAGACATCAACCAAGATGAAGATGGCAAGCTAGGATAGGAAGTTCGCG 1119
Db 398 GAGGAGTCAGGAGACACCAACGAGATGCGCATGGCAGGAGATAGGAGAGTTGCGCG 457
QY 1120 GTACAAAGACCTCTACTTGGAACTTTTGGCACACAGAAAGAGCTGCAGAGGCAATACG 1179
Db 458 GAAACAAAGACCTCTACTTGGAACTTTTGGCACACCAAGAGAGAGCTGCTGAGGCTATG 517
QY 1180 ACATTGCGGCATCAAAATTCAGAGATTAACCGAGTCACTAACTTCGACATGAACAGAT 1239
Db 518 ACATTGCTGCTATCAAAATTCAGGAGTTAAATGCAATGCAAACTTTGACATGAGTCGCT 577
QY 1240 ACAAGCTTAAAGCAATTCCTCGAAAGCCCTAGTC 1272
Db 578 ACCATGTAAGAGCATTGCAAAATAGCACTCTTC 610

RESULT 15
US-10-424-599-70525
Sequence 70525, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kowalic, David K
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: B1458 US NA
CURRENT APPLICATION NUMBER: US/10/424,599

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 19, 2004, 14:52:42 ; Search time 4152 Seconds
(without alignments)
4164.307 Million cell updates/sec

Title: US-09-980-364-2
Perfect score: 3119
Sequence: 1 MNNWLGSLSPYEQNHHRK.....YHGGGGEVAPFTVWNND 579

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2/USPFO.spool/US09980364/runat_15072004_095236_25880/app_query.fasta_1.775
-DB=EST -Qfmt=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09980364 @CGN 1.1 3437 @runat_15072004_095236_25880 -NCPU=6 -ICPU=3
-NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	992	31.8	769	14	CA783156	CA783156 sat20d05.	
2	949.5	30.4	1753	11	AY109146	AY109146 Zea mays.	
C	3	947.5	30.4	556	28	BZ512981	BZ512981 EMTQP05TF
	4	940.5	30.2	882	14	CK267021	CK267021 EST713099
	5	934.5	30.0	590	12	BM094116	BM094116 sah25f05.
	6	899.5	28.5	748	14	CF807326	CF807326 pSH023xf
7	875	28.1	720	13	BQ864461	BQ864461 QGC26M12.	
8	872.5	28.0	552	13	BU964857	BU964857 sat04e07.	
9	869.5	27.9	593	12	BM307997	BM307997 sak40e09.	
10	851.5	27.3	644	10	AW200688	AW200688 ses2c07.y	
11	849	27.2	1160	14	CK206573	CK206573 FGAS04817	
12	841	27.0	555	12	BJ188928	BJ188928 BU188928	
13	841	27.0	558	12	BJ178045	BJ178045 BU178045	
14	838	26.9	690	13	CA094556	CA094556 SCCCL400	
15	834	26.7	500	10	AW780688	AW780688 sl75e07.y	
16	820	26.3	545	14	CD475882	CD475882 nad03-15m	
17	815	26.1	631	13	CA103041	CA103041 SCEQHR108	
18	811	26.0	674	10	BF647766	BF647766 NF025G09E	
19	809	25.9	558	14	CA232734	CA232734 SCRFJL306	
20	809	25.9	599	14	CA230632	CA230632 SCRFJL306	
21	796	25.5	697	14	CA189006	CA189006 SCCCL400	
22	781.5	25.1	585	12	BM086088	BM086088 sah35d02.	
23	781.5	25.1	640	13	CA100375	CA100375 SCVPCL606	
24	743.5	23.8	641	14	CD879292	CD879292 AZO4. 104N	
25	713	22.9	420	9	AJ475492	AJ475492 AJ475492	
26	700.5	22.5	466	13	BU965263	BU965263 sat08c04.	
27	699	22.4	596	12	BM307877	BM307877 sak39b09.	
28	699	22.4	935	14	CK261980	CK261980 EST708058	
29	694	22.3	704	13	BQ625052	BQ625052 USDA-PP.0	
30	691	22.2	907	14	CK289457	CK289457 EST752179	
31	688.5	22.1	469	14	CF244329	CF244329 HDN--06-D	
C	32	688	22.1	497	14	CF244784	CF244784 3530.1.5
	33	685.5	22.0	842	14	CF446573	CF446573 EST682918
	34	676.5	21.7	513	13	BQ123372	BQ123372 EST607948
	35	675.5	21.7	545	12	B1974354	B1974354 sai99f02.
36	672	21.5	1897	11	AY103852	AY103852 Zea mays	
37	670.5	21.5	476	12	BM092890	BM092890 saj02d05.	
38	665.5	21.3	956	14	CK294077	CK294077 EST756791	
39	655.5	21.0	955	14	CK257531	CK257531 EST741168	
40	645.5	20.7	948	14	CK259208	CK259208 EST741281	
41	645	20.7	679	12	BF312281	BF312281 QCG35h06.	
C	42	640.5	20.5	728	14	CF035870	CF035870 QCG37g01.
	43	635.5	20.4	694	14	CF036846	CF036846 QCG37g01.
	44	635.5	20.4	737	14	CD832004	CD832004 EN40.061K
	45	631	20.2	527	12	BM309412	BM309412 sak58c10.

ALIGNMENTS

RESULT 1
CA783156
LOCUS
DEFINITION CA783156 769 bp mRNA linear EST 04-DEC-2002
sat20d05.v1 Gm-cl036 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl036-14361 5' similar to TR:Q9SYC2 Q9SYC2 F11M15.6 PROTEIN. ;
mRNA sequence.
CA783156
CA783156.1 GI:26045627
EST.
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE AUTHORS

1 (bases 1 to 769)
Shoemaker, R., Keim, P., Vodkin, L., Expelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, F., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.
Public Soybean EST Project
Unpublished (1999)

TITLE JOURNAL COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: cu@resgen.com web site: www.resgen.com

Seq primer: -40RP from Gibco
High quality sequence stop: 454.

FEATURES

source

1. 769
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl036-14361"
/tissue_type="somatic embryos cultured on MSD 20"
/lab_host="DH10B"
/clone_lib="Gm-cl036"
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers/adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E. coli ElectroMax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Alignment Scores:

Pred. No.: 8,04e-87 Length: 769
Score: 992.00 Matches: 187
Percent Similarity: 89.21% Conservative: 28
Best Local Similarity: 77.59% Mismatches: 14
Query Match: 31.81% Indels: 12
DB: 14 Gaps: 3

US-09-980-364-2 (1-579) x CA783156 (1-769)

QY 160 LysGlyLeuSerLeuSerMetAsnSer-----SerThrSerCysAspAsnAsnAsp 177
|||
Db 20 AATCATGACATATTCAGGAGGAGTGAAGATCAACATGGAACACGAGTGTGAA 79
|||
QY 178 SerAsnAsnValValAlaGlnGlyLysThrIleAspSerValGluAlaThrPro 197
|||
Db 80 AATAGCACAAACACT-----ACTGTTGAAATAGCACCT 112
|||
QY 198 LysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHis 217
|||
Db 113 AGAAGAACTTGGATACATTCGGGCGAGAACATCATATATCGTGAGTAACCTGCAT 172
|||
QY 218 ArgTrpThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGlyGln 237
|||

Db 173 AGATGACTGGAAGGTATCAAGCTCATCTTTGGATAATAGCTGTAGAGGAGGCCAA 232
|||
QY 238 ThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGluGlyLysAlaArg 257
|||
Db 233 TCAAGAAAAAGGACGCCAAGTTTATTGGGTGATATGATAAAGAGAGAAACACCTAGA 292
|||
QY 258 AlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyTyrThrThrThrThrAsnPheProMet 277
|||
Db 293 GCTTATGATTAGTCTCTGAAGTACTGGGGGACATCCACCACTTCCCAATT 352
|||
QY 278 SerGluTyrGluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValala 297
|||
Db 353 AGCAACTATGAGAAGGAATGGATGAAATGAAACACATGACGAGACAAGAATTTTGTGCC 412
|||
QY 298 SerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThr 317
|||
Db 413 GCATTAGAGGAAAAAGCAGTGGTTCTCCAGGGGTGCATCAATGATCGTGGAGTTACA 472
|||
QY 318 ArgHisHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAsp 337
|||
Db 473 AGGCATCACCAACACGGAAGATGGCAAGCAAGGATTGGCAGAGTTGCAGGAAACAAGAT 532
|||
QY 338 LeuTyrLeuGlyThrPheGlyThrGlnGluAlaGluAlaTyrAspIleAlaAla 357
|||
Db 533 CTTTACTTTGGGAACCTTTCAGTACTGAGGAAGAGGCTGCAGAGGATACGATAGCAGCG 592
|||
QY 358 IleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLys 377
|||
Db 593 ATAAAGTTTCAAGAGTCTCAACGCTGTCAAACTTTTGACATGAGCCGATACGAGTGA 652
|||
QY 378 AlaIleLeuGluSerProSerLeuProfile---GlySerAlaAlaLysArgLeuLysGlu 396
|||
Db 653 GCATCTTTGAAAGCAACACTCTCCCAATAGGAGGCGCTGCAAGAGCGTCTGAAAGAA 712
|||
QY 397 Ala 397
|||
Db 713 GCT 715

RESULT 2

LOCUS AY109146 1753 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays PCO137288 mRNA sequence.
ACCESSION AY109146
VERSION AY109146.1 GI:21212582
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE

1 (bases 1 to 1753)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
Unpublished (2002)
2 (bases 1 to 1753)
Coe, E.H.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

JOURNAL

Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSI, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

Location/Qualifiers
1. 1753
/organism="Zea mays"
/mol_type="mRNA"

```

/db_xref="MaizeDB:638053"
/db_xref="taxon:4577"
/cloned_lib="Maize Mapping Project/DuPont Cornsensus
Library"
/notes="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed Dupont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN
Alignment Scores:
Pred. No.: 4,6e-82 Length: 1753
Score: 949.50 Matches: 233
Percent Similarity: 48.46% Conservative: 50
Best Local Similarity: 39.90% Mismatches: 130
Query Match: 30.44% Indels: 171
DB: 11 Gaps: 18

US-09-980-364-2 (1-579) x AY109146 (1-1753)
QY 3 AsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisHisArg-----Lys 20
DB 63 CACCATTGGCTCTCCTTCCTCCTCCTC-----AACCACTACCACCATGGCTACTC 113
QY 21 AspValTyrSerThrThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
DB 114 GAAGCTTCTTAACTCTCCGGTACTCTCTT-----146
QY 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
DB 147 -----GGAGACGACGAGCGGCGCACTGGAGGAGTCC-----176
QY 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAspIleAsn 80
DB 176 -----176
QY 81 GlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLys---LeuGluAsn 99
DB 177 -----CCGAGGACGGTGGAGGAC 194
QY 100 PheLeuGlyArgThrThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGly 119
DB 195 TTCTTCGGC-----GGC 206
QY 120 SerGlyCysTyrGlyGlyAspGlyGly-----130
DB 207 GTCGGTGGTTCGGCGCCCGCCGCGCGCTGCAGATCAGGATCACCAGCTGTG 266
QY 131 ---GlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProValAspAsn 149
DB 267 TCGCGCGAGCTGGGC---AGCATCACAGCCAGGTCTTTCGCGCCACTACCCGCGCGGCCCA 323
QY 150 ValAspAsnGln---GluAsnGlyAsnAlaAlaLysGlyLeuSerLeuSerMetAsnSer 168
DB 324 GCTGGGACGACGCGTGGAGAACCCCGCGCGGTGACCGTGGCGGCATGCTGTCGACGGAC 383
QY 169 SerThrSerCysAspAsnAsnAspSerAsnAsnValValAlaGlnGlyLysThr 188
DB 384 GTGGCGGGGGCGGAGTCCGACCGAG-----407
QY 189 IleAspAspSerValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThr 208
DB 408 -----CGGAGGCGGCCCGCGAGAGCTTCGCCCGACGACCA 443
QY 209 SerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHisLeuTrp 228
DB 444 TCATCTACCGTGGCGTCCACGAGGACCGGTGGAGCGGGAGATATGAGGCGCACCTGTGG 503
QY 229 AspAsnSerCysLysArgGluGlnThrArgLysGlyArgGlnValTyrLeuGlyGly 248
DB 504 GACACACACTGCCGCGGGAGGGCCAAAGCCGCAAGGACGCCCAAGTCTACCTAGGAGGC 563

```

RESULT 3

BZ512981/c

LOCUS

BZ512981

DNA

linear

GSS 16-DEC-2002

BOMQP05TF BO_2_3_KB Brassica oleracea genomic clone BOMQP05,

Genomic survey sequence.
 BZ512981
 BZ512981.1 GI:27039560
 GSS.
 Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 556)
 Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
 Whole genome shotgun sequencing of *Brassica oleracea*
 Unpublished (2001)
 Other GSSs: BOMQPO5TR
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TF
 Class: sheared ends.

FEATURES

source
 1..556
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="Tol000DH3"
 /db_xref="taxon:3712"
 /clone="BOMQPO5"
 /clone_lib="BO_2_3_KB"
 /note="Vector: pHO51; Site 1: BstXI; 2-3 kb sheared
 Genomic DNA inserted into pHO51 using BstXI linkers"

ORIGIN

Alignment Scores:
 Pred. No.: 1,14e-82 Length: 556
 Score: 947.50 Matches: 183
 Percent Similarity: 98.39% Conservative: 0
 Best local Similarity: 98.39% Mismatches: 3
 Query Match: 30.38% Indels: 2
 DB: 28 Gaps: 0

US-09-980-364-2 (1-579) x BZ512981 (1-556)

QY 377 LysalalleLeuGluSerProSerLeuProleGlySerAlaAlaLysArgLeuLysGlu 396
 DB 556 AAGCATCCCTCGAAGCCCTAGTCT-CCTATTGT-AGCGCGCAAAACGTTCAAGGAG 499
 QY 397 AlaAsnArgProValProSerMetMetMetIleSerAsnAsnValSerGluSerGluAsn 416
 DB 498 GCTAACCGTCGGTTCGAAGTATGATGATCATGATCATGATTAACGTTTCAGAGAGTGAGAAT 439
 QY 417 SerAlaSerGlyTrpGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeu 436
 DB 438 AGTGCTAGCGGTGGCAAAACGCTGCGGTCAGCATCATCAGGAGTAGATTGACCTTA 379
 QY 437 LeuHisGlnHisGlnGluArgTyrAsnGlyTyrTrpAsnGlyGlyAsnLeuSerSer 456
 DB 378 TTGCACCAACATCAAGAGAGTCAATGGTTATATTATACATCGAGAAACCTTGTCTTCG 319
 QY 457 GluSerAlaArgAlaCysPheLysGlnGluAspAspGlnHisPheLeuSerAsnThr 476
 DB 318 GAGAGTCTAGGCTTCTTCAACAGAGAGTATGATCAACACCATTTCTTGACCAACAG 259
 QY 477 GlnSerLeuMetThrAsnIleAspHisGlnSerSerValSerAspSerValThrVal 496
 DB 258 CAGAGCCTCATGACTATATATCATCATCAAGTCTGTTTCGGATGATTCGGTTACTGTT 199
 QY 497 CysGlyAsnValValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAsp 516
 DB 198 TGTGGAAATGTTGTGGTATGTGGTTATCAAGGATTTGCAGCCCCGGTTAACTCGGAT 139

QY 517 AlaTyrAlaAlaSerGluPheAspTyrAsnAlaAaArgAsnHisTyrTyrPheAlaGlnGln 536
 DB 138 GCCTACGCTGCTAGTGGATTGATTATACGCAAGAAACCATTTACTTTGCTCAGCAG 79
 QY 537 GlnGlnThrGlnGlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGly 556
 DB 78 CAGCAGACCAGCAGTCGCCAGGTGGAGATTTTCCCGCGCAATGACGAATAATGTTGC 19
 QY 557 SerAsnMetTyrTyrHis 562
 DB 18 TCTAATATGTTATTACCAT 1

RESULT 4
 CK267021
 LOCUS
 DEFINITION
 CK267021 potato abiotic stress cDNA library Solanum tuberosum cDNA
 clone POAC822 5' end, mRNA sequence.
 CK267021
 CK267021.1 GI:39823999
 EST.
 Solanum tuberosum (potato)
 Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 882)
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
 Unpublished (2003)
 Other ESTs: EST713100
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source
 1..882
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="POAC822"
 /tissue type="abiotic stress treated leaf and root tissue"
 /lab host="DH10B-Tona"
 /clone_lib="potato abiotic stress cDNA library"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: Solanum tuberosum var. Kennebec plants were
 grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to
 four separate sets of plants. Set 1 involved saturation of
 the soil with 150 mM NaCl and tissues were harvested at
 following application of the salt stress (leaves: 2hr,
 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1 d,
 Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering
 applications. Drought stressed plants were harvested after
 cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
 and 5d). Set 3 were grown under the standard conditions
 and then were cold stressed by placement at 4 C. Cold
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d. Set 4 were grown under the standard conditions and
 then were heat stressed by placement at 35 C. Heat
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d and 4d and heat-stressed roots were harvested at 6 hr,
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and
 equal RNA from each tissue and stress was pooled to
 construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:
 Pred. No.: 1.17e-81 Length: 882

```
Score: 940.50 Matches: 191
Percent Similarity: 73.38% Conservative: 35
Best Local Similarity: 62.01% Mismatches: 50
Query Match: 30.15% Indels: 50
DB: 14 Gaps: 7

US-09-980-364-2 (1-579) x CK257021 (1-882)

QY 156 GlyAsnAlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsn 175
Db 43 GCGCGGGAGGTGGAGCTTTGCTTAGCAGTTAATGCTGACGACAACTACTACTACAG 102
QY 176 AsnAspSerAsnAsnValValAlaGlnGlyLysThrIleAspAspSerValGluAla 195
Db 103 TGTCTCGAAATGAAAGGCGATGTT-----GCTGTTCTGATTCTCAGAGTTGT 153
QY 196 ThrProLysThrIleGlnSerPheGlyGlnArgThrSerIleTyArgGlyValThr 215
Db 154 -----AAGAAATTTCTGATACTTTTGGCAAGAACTTCAATCTACAGAGGTGTAA 207
QY 216 ArgHisArgTrpThrGlyArgTyGluAlaHisLeuTrpAspAsnSerCysLysArgGlu 235
Db 208 AGACATAGATGGACAGAAATATGAAGCTCATCTATGGATTAACAGCTGTAGGAGAA 267
QY 236 GlyGlnThrArgLysGlyArgGlnValTyLeuGlyLysThrAspLysGluLysAla 255
Db 268 GGCACAGCTAGAAAGGGCGTCAAGTGTACTTGGGTGATGATACAAAGAAATAAGCA 327
QY 256 AlaArgAlaTyArgLeuAlaLeuLysTyTrpGlyThrThrThrThrAsnPhe 275
Db 328 GCAAGGGCATATGACTTGGCAGCTCTAAGTATTTGGGTCCACAGCTACCACTTC 387
QY 276 ProMetSerGluTyGluLysGluValGluLysMetLysHisMetThrArgGlnGluTy 295
Db 388 CTGTACTATATATACTAAGAAATTCGAGGAATGAACACATGACTAAGCAAGATTC 447
QY 296 ValAlaSerLeuArgLysSerSerGlyPheSerArgGlyAlaSerIleTyArgGly 315
Db 448 ATTGCTCTCTAGAGAGAAAGTAGTGGTTCTCCGAGGAGCTTCGATTACCGGGT 507
QY 316 ValThrArgHisHisGlnHisGlyArgTrpGlnAlaArgLysGlyArgValAlaGlyAsn 335
Db 508 GTGACAGGCGCATCATCAACAGGCGCTGGCAACGAGAAATGGCCAGTTGTGGAAAC 567
QY 336 LysAspLeuTyLeuGlyThrPheGlyThrGlnGluAlaAlaGluAlaTyArgPhe 355
Db 568 AAGGATCTTTACCTTGGAAACATTTGCTACTGAGGAGGAAGCAGCGAAGCGTATGAC 627
QY 356 AlaAlaLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyAsn 375
Db 628 GCAGCCATAAGTTTCAGGGGAGTGAATGCAGTGACCACTTCGAGATCAATCGTTATG 687
QY 376 ValLysAlaLysLeuGluSerProSerLeuProIleGlySerAlaAlaLysArgLeu 395
Db 688 CTCGAGGCATCATGCAAGT---TCCCTCCAGTTGGTGGAAACCGCTAAGCGTTTGA 744
QY 396 GluAlaAsnArgProValProSerMetMetMetIleSerAsnAsnValSerGluSer 415
Db 745 CTCCTCC-----CTTGAATCCGAG 762
QY 416 AsnSerAlaSerGlyTrpGlnAsnAlaAlaValGlnHisHisGlnGlyValAspLeu 435
Db 763 CAGAATCATCTCCAGCAATAAC-----AACACACCACGAGCAGCAGCTTGACT 813
QY 436 LeuLeuHisGlnHisGlnGluArgTyArgGlnGlyTyTrpAsnGlyGlyAsnLeuSer 455
Db 814 CA-ATGCAACAGCACCAACAA-----CATCAATTTCGGGGC----- 848
QY 456 SerGluSerAlaArgAlaCysPhe 463
Db 849 ----AATGCCACCGCTCTGCTAT 869

RESULT 5
```

QY 223 TyrGluAlaHisLeuTrpAspAsnSerCysLysArgGluGlnThrArgLysGlyArg 242
 Db 61 TATGAAGCTCACCTTTGGGATAATAGCTGTAGAAGGGAGGCAATCAAGAAGAAGGAGC 120
 QY 243 GlnValTrpLeuGlyGlyTrpAspLysGluGlnLysAlaAlaArgAlaTrpAspLeuAla 262
 Db 121 CAAAGTTATTTGGTGGATATGATAAAGAGAAAGACGAGCTAGTCTTATGATTAGCT 180
 QY 263 AlaLeuLysTrpGlyTrpGlyThrThrThrThrAsnPheProMetSerGluTrpGluLys 282
 Db 181 GCACTCAAGTACTGGGGACATCCACCACTACCACTTTCCAATTAGTAACATGAGAG 240
 QY 283 GluValGluGluMetLysHisMetThrArgGlnGluTrpValAlaSerLeuArgLys 302
 Db 241 GAATTGGATGAATGAACACATGACGCGACAGAAATTTGTTGCTGCCATTAGAAGAAA 300
 QY 303 SerSerGlyPheSerArgGlyAlaSerIleTrpArgGlyValThrArgHisGlnHis 322
 Db 301 AGCAGTGGTTTCTCCAGGGGTGCATCAATGATCTGAGTTACAGGCATCACCACAC 360
 QY 323 GlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTrpLeuGlyThr 342
 Db 361 GGAAGATGGCAAGCAAGAAATTTGGCAGAGTTGCGAGAAACAAAGATCTTTACTTGGAACT 420
 QY 343 PheGlyThrGlnGluAlaAlaGluAlaTrpAspIleAlaAlaIleLysPheArgGly 362
 Db 421 TTCAGTACTGAGAGAGAGGCTGCTGAGCATACGATAGCTGCATTAAGTTCAGAGGT 480
 QY 363 LeuThrAlaValThrAsnPheAspMetAsnArgTrpAsnValLysAlaIleLeuGluSer 382
 Db 481 CTCAACGGTGTCAACAACTTTGACATGAGCGCTACGACGTGAAAGCCATCTCTTGAAGC 540
 QY 383 ProSerLeuProIle---GlySerAlaAlaLysArgLeuLysGluAla 397
 Db 541 AACACTCTCCCATAGGAGGAGGAGCTGCAAGGCGTCTGAAGAAGCT 588

RESULT 6
 CF807326 748 bp mRNA linear EST 27-OCT-2003
 LOCUS psHB023xF07f USDA-IFAFS:Expression of Phytophthora sojae genes
 DEFINITION during infection and propagation of Phytophthora sojae cDNA clone
 SH023F07 5, mRNA sequence.

ACCESSION CF807326
 VERSION 1
 KEYWORDS EST
 SOURCE Phytophthora sojae
 ORGANISM Phytophthora sojae

REFERENCE Tyler B.
 AUTHORS Tyler B.
 TITLE Not Published
 JOURNAL Unpublished (2003)
 COMMENT Contact: Tyler B
 Tyler lab

Phytophthora

1 (bases 1 to 748)

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

/cell_line="P6497"
 /dev_stage="48 hr. post infection stage"
 /lab_host="Soybean plant"
 /clone_lib="USDA-IFAFS:Expression of Phytophthora sojae
 genes during infection and propagation"
 /notes="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Alignment Scores:
 Pred. No.: 9,31e-77 Length: 748
 Score: 889.50 Matches: 174
 Percent Similarity: 81.51% Conservative: 20
 Best Local Similarity: 73.11% Mismatches: 19
 Query Match: 28.52% Indels: 25
 DB: 14 Gaps: 4
 US-09-980-364-2 (1-579) x CF807326 (1-748)
 QY 161 GlyLeuSerLeuSerMetAsnSerSerThrSerCys----- 172
 Db 39 GCGCGTTCTCTCCCAACAGCCACCGCCACCGTGTGCACCGCCACAGCTTCCGAGTTC 98
 QY 173 -----AspAsnAsnAsn-----AspSerAsn----- 179
 Db 99 TCCACCGCACCAACCAACCCACCTGTACGATTCCGAGCTGAAGACAAATAGCCCGGTGC 158
 QY 180 -----AsnAsnValValAlaGlnGlyLysThrIleAspAspSerValGluAlaThrPro 197
 Db 159 TTCCTCTCGCGCTTTGTGCGCGAACCACCAACC-----GAACCTCAGAAACCCCTCTCCA 212
 QY 198 LysLysThrIleGluSerPheGlyGlnArgThrSerIleTrpArgGlyValThrArgHis 217
 Db 213 AAAAAACCCGTGCACACCTTCGCGCAACGACCTCCATCTACCGCGGCTCACCGGACAT 272
 QY 218 A:G:TrpThrGlyArgTrpGluAlaHisLeuTrpAspAsnSerCysLysArgGluGln 237
 Db 273 AGATGGACGGGAAGATACGAAGCTCATCTATGGGACATAGTTGTAGAAGAGAGGCCAA 332
 QY 238 ThrArgLysGlyArgGlnValTrpLeuGlyGlyTrpAspLysGluGluLysAlaAlaArg 257
 Db 333 ACAGAGAAAGGAGGACAAAGTTTACCTGGTGGTTATGACAAGGAGATAGAGGAGCCAGG 392
 QY 258 AlaTrpAspLeuAlaAlaLeuLysTrpGlyThrThrThrThrThrThrThrThrThr 277
 Db 393 GCTTACGATCTCGAGCTCTCAAGTACTGGGTCCAACTACCACTACCACTTTCCCAAT 452
 QY 278 SerGluTrpGluLysGluValGluMetLysHisMetThrArgGlnGluTrpValAla 297
 Db 453 TCCAACTATGAGAGAACTGGAGGAGATGAAGAACATGACGAGGAGAGATTGTTGCT 512
 QY 298 SerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTrpArgGlyValThr 317
 Db 513 TCTCTACGAGGAGAGAGCGAGTGGTTTCTTAGGGGGCTCTATATACAGAGGAGTACG 572
 QY 318 ArgHisHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAsp 337
 Db 573 AGACACCAACGAGCATGGCGGATGGCAGCGAGATAGGAGAGTTGCCGGAACAAAGAC 632
 QY 338 LeuTrpLeuGlyThrPheGlyThrGlnGluAlaAlaGluAlaTrpAspIleAlaAla 357
 Db 633 CTCTACCTTGGAACTTTTCCAGCAACAAAGAAAGCTGCTGAGGCTATGACATTGCTGCT 692
 QY 358 IleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTrpAsn 375
 Db 693 ATCANATTACGGGATTAATGTCAGTCAGTCANCAACTTTGACATGCTCCTACGAT 746

RESULT 7

BQ864461

LOCUS

DEFINITION

ACCESSION

VERSION

1.748

Phytophthora sojae

1880 Pratt Dr., Blacksburg, VA 24061, USA

Tel: 540-231-7318

Email: bmtaylor@vt.edu

PCR Primers

FORWARD: BK reverse primer

BACKWARD: BK reverse primer

Plate: 023 row: F column: 07

Seq primer: BK reverse primer

High quality sequence stop: 748.

Location/Qualifiers

1.748

/organism="Phytophthora sojae"

/mol_type="mRNA"

/db_xref="taxon:67593"

/clone="SH023F07"

/tissue_type="mycelium"

source

Phytophthora sojae

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

Phytophthora

```

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
    Koziak,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
    Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J.,
    Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y.,
    Lai,Z., Church,S., Jackson,L. and Bradford,K.
    Lettuce and Sunflower ESTs from the Compositae Genome Project
    http://comphenomics.ucdavis.edu/
    Unpublished (2002)
    Contact: Alexander Koziak [R.W.Michelmore]
    Department of Vegetable Crops, R.W.Michelmore Lab
    University of California at Davis (UCD)
    Asmundson Hall, UCD, Davis, CA 95616, USA
    Tel: 1-(530)-742-1742
    Fax: 1-(530)-752-9659
    Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
    singleton, see http://cgdb.ucdavis.edu/ for details.
    Plate: QGC26 row: M column: 12.

FEATURES
    source
    1..720
        /organism="Lactuca sativa"
        /mol_type="mRNA"
        /cultivar="Salinas"
        /db_xref="taxon:4236"
        /clone="QGC26M12"
        /lab_host="E.coli"
        /clone_lib="QGC-ABCDI lettuce salinas"
        /notes="Vector: pBRCN4SfiAB; The library was constructed
        from 10 different sources of RNA from a single genotype.
        Separate cDNAs were generated using primers that
        incorporated unique 5' and 3' tags to distinguish each
        source of RNA. cDNAs were then pooled, size-fractionated,
        directionally cloned into a custom medium-copy vector and
        transformations made with four size classes to minimize
        size bias. Details of each source of RNA and library
        construction can be obtained at http://cgdb.ucdavis.edu/
        TAG_TISSUE=chemical induction
        TAG_LIB=QGC ABCDI lettuce salinas
        TAG_SEQ=TGTAGCGGG"

ORIGIN
Alignment Scores:
Pred. No.: 2.34e-75 Length: 720
Score: 875.00 Matches: 171
Percent Similarity: 81.93% Conservative: 24
Best Local Similarity: 71.85% Mismatches: 32
Query Match: 28.05% Indels: 12
DB: 13 Gaps: 2

US-09-980-364-2 (1-579) x B0864461 (1-720)

Qy 222 ArgTyrGluAlaHisLeuTyrAspAsnSerCysLysArgGluGlyGlnThrArgLysGly 241
Db 2 AGATACGAGGCTCATTTATGCGACAAATAGTTGCAGAGAGAGGACCAAGTTCGCAAGGA 61
Qy 242 ArgGlnValTyrLeuGlyGlyTyrAspLysGluLysAlaAlaArgAlaTyrAspLeu 261
Db 62 AGACAAGTTTATTTGGGTGGTTACGACAAAGAGATAGGCGAGAGGCTTATGATTTA 121
Qy 262 AlaAlaLeuLysTyrTyrGlyThrThrThrThrThrThrThrThrThrThrThrThrThr 281
Db 122 CTGTGATTGAATATTTGGGTGGGTACCAACTACGACAAATTTCCCTATTAGCAACTATGAG 181
Qy 282 LysGluValGluGluMetLysHisMetThrArgGlnGlnTyrValAlaSerLeuArgArg 301
Db 182 AAGGAGATCGGAGGATGAACACATGACTAGACAAAGATACGTAGCATCTATTAGGAGG 241

```

```

Qy 302 LysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHisGln 321
Db 242 AAAAGTAGCGGTTTTCTCGAGGTGCTTCATCTATAGAGGTGTGACAGACACCATCAA 301
Qy 322 HisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGly 341
Db 302 CATGGAAGATGGCAAGCAAGAAATTGGAAGAGTTGCAGGAACAAGATCTTTACTCGGA 361
Qy 342 ThrPheGlyThrGlnGlnGluAlaAlaGluAlaTyrAspIleAlaIleLysPheArg 361
Db 362 ACTTTTGGTACACAGGAAGAGCTGCTGAGGCTTATGACATTGCTGCATCAATTCGCT 421
Qy 362 GlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGlu 381
Db 422 GGACTGATGTCAGTCACAAATTTGAATCAACAGATACGACGTTAAATGCATACTTGA 481
Qy 382 SerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaA-SenArgProVa 401
Db 482 AGCACCACTTACCGGTTGGTGGTGCAGCTAAACGCTCTCAAGGATGCGAACACGCGCT 541
Qy 401 lpro-----SerMetMetMetIleSerAsnAsnValSerGluSe 414
Db 542 GCCACTGATGTTACTTCGGGCTGCAAAAGATGATCAC-TTGACCACCACTGACCCACGA 600
Qy 414 rGluAsnSerAlaSerGly-----TrpGlnAsnAlaAlaValGlnHisHisGlnGln 431
Db 601 GGGACTCAACAGTTATGTCACCATGACATGCGCGAGCTCCATTAACTACCAACCAAGC 660
Qy 431 yValAspLeuSerLeuLeuHisGlnHisGlnGluArgTyrAsnGlyTyrTyr 448
Db 661 CTTGCTCAAGCACAGGCGCAACCGTTTAGCATGCATATCCACACACTAT 712

RESULT 8
BU964897
LOCUS
DEFINITION
    sac04a07.yl Cm-cl036 Glycine max cDNA clone SOYBEAN CLONE ID:
    Cm-cl036-12685 5' similar to TR:Q9SYC2 Q9SYC2 F11M15.6 PROTEIN. ;
    mRNA sequence.
ACCESSION
    BU964897
VERSION
    BU964897.1 GI:24205644
KEYWORDS
    EST.
SOURCE
    Glycine max (soybean)
ORGANISM
    Glycine max
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
    Glycine.
REFERENCE
    1 (bases 1 to 552)
    Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
    Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
    Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
    Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
    Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
    McCann,R., Waterston,R. and Wilson,R.
    Public Soybean EST Project
    Unpublished (1999)
    Contact: Shoemaker R/Public Soybean EST Project
    Public Soybean EST Project
    Washington University School of Medicine
    4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
    Tel: 314 286 1800
    Fax: 314 286 1810
    Email: est@wustl.edu
    This clone is available through: ResGen, Invitrogen Corp. 2130
    South Memorial Parkway Huntsville, AL 35801 For further information
    call: (800)-533-4363 or contact: ccu@resgen.com web site:
    www.resgen.com
    Seq primer: -40RP from Gibco
    High quality sequence stop: 445.
FEATURES
    source
    1..552
        /organism="Glycine max"
        /mol_type="mRNA"

```

/db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl036-12685"
 /tissue_type="somatic embryos cultured on MSD 20"
 /lab_host="DH10B"
 /clone_lib="Gm-cl036"
 /note="Vector: pSPOR1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPOR1 vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

ORIGIN

Alignment Scores:
 Pred. No.: 2,686-75 Length: 552
 Score: 872.50 Matches: 162
 Percent Similarity: 96.72% Conservative: 15
 Best Local Similarity: 88.52% Mismatches: 5
 Query Match: 27.97% Indels: 1
 DB: 13 Gaps: 1

US-09-980-364-2 (1-579) x BU964897 (1-552)

QY 213 GlyValThrArgHisArgTyrGlyArgTyrGluAlaHisLeuTyrAspAsnSerCys 232
 DB 3 GGAGTAACGGACATAGATGACTGGAGGTATGAGCTACCTTTGGGATATAGCTGT 62
 QY 233 LysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGlu 252
 DB 63 AGAAGGGAAGGCGAATCAAGAAAGGAGCCCAAGTTTATTGGGTGATATATAAGAA 122
 QY 253 GluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyThrThrThr 272
 DB 123 GAATAAGCAGCTAGTCTTATGCTTACCTGAGTACTGGGGGACATCCACCACT 182
 QY 273 ThrAsnProMetSerGluTyrGluLysGluValGluGluMetLysHisMetThrArg 292
 DB 183 ACCAACTTTCCAAATAGTAAGTATGAGAGGAATGGATGAATGAACACATGACGCGA 242
 QY 293 GlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIle 312
 DB 243 CAAGAATTGTGTGTCATTAGAAAGAAAGACAGTGTCTTCCAGGGGTGCATCAATG 302
 QY 313 TyrArgGlyValThrArgHisHisGlnHisGlnValArgTyrGlnAlaArgIleGlyVal 332
 DB 303 TATCGTGGAGTTACAGGCATCACACACGAGATGGCAAGCAAGAAATTCACAGATT 362
 QY 333 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluAlaAlaGluAla 352
 DB 363 GCAGGAACAAGATCTTTACTTGGGAACCTTCAGTACTGAAGAGAGGGTGTGTAAGCA 422
 QY 353 TyrAspIleAlaAlaLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsn 372
 DB 423 TACGACATAGCTGCATTAAGATTTCAGAGGTCTCAACGCTGTCTCAAACTTTGATGAGC 482
 QY 373 ArgTyrAsnValLysAlaLysLeuGluSerProSerLeuProIle---GlySerAlaAla 391
 DB 483 CGCTACACGCTGAAGCATCTTCTTGAAGCAACACTCTCCCAATAGGAGGAGCTGCA 542
 QY 392 LysArgLeu 394
 DB 543 AAGGCTCTG 551

RESULT 9

BM307997
 LOCUS
 DEFINITION
 sak40e09.y1 Gm-cl036 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl036-5586 5' similar to TR:Q95YC2 Q95YC2 FILM15.6 PROTEIN. ; mRNA sequence.
 BM307997
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Glycine max (soybean)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 REFERENCE
 1 (bases 1 to 593)
 Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,F., Gibbons,M., Fape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project
 Unpublished (1999)
 CONTACT: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 427.
 Location/Qualifiers
 1..593
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl036-5586"
 /tissue_type="somatic embryos cultured on MSD 20"
 /lab_host="DH10B"
 /clone_lib="Gm-cl036"
 /note="Vector: pSPOR1; Site 1: NotI; Site 2: SalI; This cDNA library was constructed from mRNA isolated from somatic embryos (age ranging from 2 months to 9 months) cultured on MSD 20. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a NotI restriction site. SalI linkers adapters were ligated to the blunt-ended cDNA fragments followed by NotI digestion. The cDNA fragments were directionally cloned into the NotI-SalI restriction site of the pSPOR1 vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. This library was constructed in the laboratory of Dr. Lila Vodkin by Anu Khanna at the University of Illinois at Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

FEATURES
source

ORIGIN
 Alignment Scores:
 Pred. No.: 5,94e-75 Length: 593
 Score: 869.50 Matches: 157
 Percent Similarity: 90.10% Conservative: 25
 Best Local Similarity: 77.72% Mismatches: 15
 Query Match: 27.88% Indels: 5
 DB: 12 Gaps: 1

US-09-980-364-2 (1-579) x BM307997 (1-593)

QY 162 LeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSerAspAsnAsn 181


```

Db      654  AAGGAGCTGGAGAGATGAAGTCGATGACGGCGGAGGAGTTCATCGCTCGCTTGTAG 713
QY      302  LysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisSgln 321
Db      714  AAGAGCAGCGGCTTCTCGGAGGGCATCCATCTACAGAGGAGTAACAGGCGATCATCAG 773
QY      322  HisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGly 341
Db      774  CACGGCGGTGGCAGCAGAGATCGCAGGGTGGCCGGAAACAG-GACCTGACCTGGGA 832
QY      342  ThrPheGlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArg 361
Db      833  ACTTTTCAGCAGCAGGAGGAGCGGGAGGGGTACGACATTCGCGCGCATCAAGTTCGCG 892
QY      362  GlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGlu 381
Db      893  GGCTCACGCG-GTGACCAACTC-GACATGAG-CGCTACGAGTCGAGGAGTCCTCACA 949
QY      382  SerProSerLeuProIleGlySerAlaAlaLys 392
Db      950  GCGACCTGCCATCGCGCGGGCGGGCGCGCGCGG 982

```

BUI188928 555 bp mRNA linear EST 17-OCT-2003
 BUI188928 normalized full length cDNA library, chloronemata,
 caulonemata and malformed buds Physcomitrella patens subsp. patens
 cDNA clone pphb45d02 5', mRNA sequence.

BUI188928
 BUI188928.1 GI:18356869
 EST.

Physcomitrella patens subsp. patens
 Physcomitrella patens subsp. patens

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.

1 (bases 1 to 555)
 Nishiyama, T., Fujita, T., Shin-i, T., Seki, M., Nishide, H.,
 Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K.,
 Kohara, Y., and Hasebe, M.

Comparative genomics of Physcomitrella patens gametophytic
 transcriptome and Arabidopsis thaliana: implication for land plant
 evolution

Proc. Natl. Acad. Sci. U.S.A. 100 (13), 8007-8012 (2003)

22709184

12806149

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp

A backbone of the vector is basically from pBluescript II (KS),
 that was in vivo excised from a l-FLC phage vector (Carninci et al.
 2001). 5' end of the cDNA that was digested with XhoI was ligated
 to SalI site of the vector and the 3' end including polyA tail was
 ligated to BamHI site of the

vector(5'-GAGAGAGAGAGATCCACCTGGAGAGTTTTTTTTTTTTT-3' was
 used as a 1st 3' primer, and

5'-GGTTCGAGTCATCGCTGTCAGACAGCGATGACTCGAGAACCGNNNN-3' as 2nd

5'-hairpin primer, giving the following 5' boarder sequence,
 AGCCAAATCGCGAGCTGGAATTCGTCAGAACCG). cDNA instert could be

amplified with conventional T7 and T3 primers. This full-length
 cDNA library was generated according to the method described in

Nishiyama et al. (2003).
 Protonemata were blended by the POLYTRON, and then cultivated on

the BCDATG medium for 13-14 days under the continuous light.
 These clones are available from RIKEN Bio Resource Center

(http://www.brc.riken.go.jp/lab/epd/Eng/index.html). The database
 of Physcomitrella EST clones is available at the PHYSCOBASE

(http://moss.nibb.ac.jp):

Location/Qualifiers

1. 555
 /organism="Physcomitrella patens subsp. patens"
 /mol_type="mRNA"
 /sub_species="patens"
 /db_xref="taxon:145481"
 /clone="pphb45d02"
 /tissue_type="mixture of chloronemata, caulonemata and
 malformed buds"
 /clone_lib="normalized full length cDNA library,
 chloronemata, caulonemata and malformed buds"

ORIGIN

Alignment Scores: Length: 555
 Pred. No.: 3.4e-72 Matches: 150
 Score: 841.00
 Percent Similarity: 94.57% Conservative: 24
 Best Local Similarity: 81.52% Mismatches: 10
 Query Match: 26.96% Indels: 0
 Ds: 12 Gaps: 0

US-09-980-364-2 (1-579) x BUI188928 (1-555)

QY 191 AspSerValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIle 210
 Db 4 GACAACAAAGATCCGTCCTCCCGCCCAAGTCCATCGATACCTTCGGCCAGAGAACATCTGTG 63

QY 211 TyrArgGlyValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTyrAspAsn 230
 Db 64 TACAGAGCGGTCACTAGGCATCGTTGGACTGGAGCGGTACGAGGCACATCTCTGGACAAAT 123

QY 231 SerCysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAsp 250
 Db 124 ACTTGTAGAAAGGAAGTCAACTCGCAAGGTGCACAAGTATATCTTAGGAGGATATGAT 183

QY 251 LysGluGluLysAlaAlaArgAlaTyrAspLeuAlaLeuLysTyrTyrGlyThrThr 270
 Db 184 AAAGAAGCAAGGCGACGAGCCCTATCACTTGGCAGCGCTCAAAATATTTGGGTGTACG 243

QY 271 ThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGluMetLysHisMet 290
 Db 244 ACCACCATCACTTCACCTTGATACATACGACAAAGAGCTGGAGAGATGAAGAACATG 303

QY 291 ThrArgGlnGluTyrValAlaSerLeuArgLysSerSerGlyPheSerArgGlyAla 310
 Db 304 TCTCGCAGGAGTAGTTGCTCTGCTGTGAAAGGAAAGCAGTGGCTTTTCGAGAGAGCA 363

QY 311 SerIleTyrArgGlyValThrArgHisSglnHisGlyArgTyrGlnAlaArgIleGly 330
 Db 364 TCCATGTATCGAGGGGTGACGAGACCCATCAGCATGGAGATGCGCAGCAGCATTCGG 423

QY 331 ArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluAlaAla 350
 Db 424 CGAGTTGCGGGAAACAAGGACCTGTACCTGGGCACCTACAGCACTCAAGAGAGAGCAG 483

QY 351 GluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaValThrAsnPheAsp 370
 Db 484 GAGGCTACGACATAGCTCCATAAAATATCGTGGATTAATCCCGTCACAAACATTTTCA 543

QY 371 MetAsnArgTyr 374
 Db 544 ATCTCTCGTTAC 555

RESULT 13

BUI178045

LOCUS

DEFINITION

BUI178045

558 bp mRNA linear EST 16-OCT-2003

BUI178045

558 bp mRNA linear EST 16-OCT-2003

BUI178045

558 bp mRNA linear EST 16-OCT-2003

BUI178045

558 bp mRNA linear EST 16-OCT-2003

BUI178045

558 bp mRNA linear EST 16-OCT-2003

BUI178045

558 bp mRNA linear EST 16-OCT-2003

BUI178045

558 bp mRNA linear EST 16-OCT-2003

BUI178045

558 bp mRNA linear EST 16-OCT-2003

BUI178045

40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucrest.lad.ic.unimelb.edu.au/public>

ORIGIN

Alignment Scores: 9.52e-72 Length: 690
Pred. No.: 838.00 Matches: 169
Score: 81.36% Conservativity: 23
Best Local Similarity: 71.61% Mismatches: 26
Query Match: 26.87% Indels: 19
DB: 13 Gaps: 5

US-09-980-364-2 (1-579) x CA094556 (1-690)

QY 129 GlyGlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProValAsp 148
DB 21 GGGCGAAGCTTCATCGGGCTGTCATGATCAAGAACTGGCTGGCAGTCAGCGCGCCT 80
QY 149 AsnValAspAsnGlnGluAsnGlyAsnAlaAlaLysGlyLeuSerLeuSerMetAsnSer 168
DB 81 -----GCACCTCGGGGGTGCAC-----TCCATGGCGCTG 110
QY 169 SerThrSerCysAspAsnAsnAspSerAsnAsnValValAlaGlnGlyLysThr 188
DB 111 GCGACACGCGC---ACATCGCTGAGAGGAGGAGGAGTGTGCACCGCGGGAGAGC 167
QY 189 -----lleAspAspSerValGluAlaThrProLysLysThrIleGluSer 203
DB 168 GTTGGCATCATCTTGTGGACATCGGCAGCAGG-----AAGCCTCGGTGGACACA 221
QY 204 PheGlyGlnArgThrSerIleLysArgGlyValThrArgHisArgTrpThrGlyArgTyr 223
DB 222 TTGGGCGAGCGACTTCTATTATTCGGGGGTCCACAAAGCATAGATGGACAGAGGTAT 281
QY 224 GluAlaHisLeuTrpAspAsnSerCysLysArgGlyGlnThrArgLysGlyArgGln 243
DB 282 GAGCCCATCTTTGGGCAATAGCTGCGAAGAGAGTGTGAGTCTGCAAGCAGCAGCAA 341
QY 244 ValTrpLeuGlyGlyTrpAspLysGluGluLysAlaAlaArgAlaTrpAspLeuAla 263
DB 342 GTATACCTAGTGGATATGATATAAGAGAGAGAGCTCCAGGGCTTACGATTTAGCTGCT 401
QY 264 LeuLysTrpTrpGlyThrThrThrThrAspPheProMetSerGluTrpGluLysGlu 283
DB 402 CTCAAGTATTGGGCGACTACACCACTACAAATTTCCATGAGCAACTATGAAGAAGAA 461
QY 284 ValGluGluMetLysHisMetThrArgGlnGluTrpValAlaSerLeuArgArgLysSer 303
DB 462 CTGGAAGAGATCAAGCATATGTTCACGACAGAAATATGTCGCTGTAGAGGAAAGC 521
QY 304 SerGlyPheSerArgGlyAlaSerIleLysArgGlyValThrArgHisGlnHisGly 323
DB 522 AGTGGATTTCTCGTGGTCACTCATGATTACCGAGGGTTACCAGGCCACCCAGCATGGA 581
QY 324 ArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTrpLeuGlyThrPhe 343
DB 592 GG-TGGCAACGAGATAGAGAGAGTGGCAGAAACACAGGATCTGTTATTTAGGCACCATC 640
QY 344 GlyThrGlnGluAlaAlaGluAlaTrpAspIleAlaAlaLys 359
DB 641 AGTACCAGGAAGAGCGGAGGCTTATGACATCGTGTGCTCAAA 688

RESULT 15
AW780688 500 bp mRNA linear EST 03-DEC-2001
LOCUS sl75e07.y1 Gm-cl027 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION Gm-cl027-7165 5' similar to TR:Q41832 Q41832 OPEN READING FRAME. ?;
mRNA sequence.
ACCESSION AW780688
VERSION AW780688.1 GI:7795291
KEYWORDS EST.

SOURCE ORGANISM

Glycine max (soybean)
Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE AUTHORS

1 (bases 1 to 500)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.

TITLE JOURNAL COMMENT

Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 412.

FEATURES source

Location/Qualifiers
1..500
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl027-7165"
/tissue_type="cotyledons of 3- and 7-day-old Williams
seedlings"
/lab_host="DH10B"
/clone_lib="Gm-cl027"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from cotyledons of 3- and 7-day-old Williams seedlings
which were propagated on paper towels with distilled
water. The cotyledons were flash-frozen in liquid
nitrogen, then lyophilized for 72 hours. Unequal amounts
of mRNA was used for cDNA synthesis. Stratagene's cDNA
Synthesis Kit (catalog number 200401) was used to
synthesize the cDNA. First-strand synthesis was
performed with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An anchor
nucleotide (V=A, C, or G) was added to the 3' end of the
primer (GAGAGAGAGAGAGAGAGAGAGTCTCGAG(T)18) to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in
with cloned Pfu DNA, ligated to EcoRI adapters and
subsequently phosphorylated. The XhoI site within the
first-strand synthesis primer was then restricted by
digestion with XhoI; all XhoI sites in the cDNA would be
protected by their hemimethylated status. The cDNA
constructs were size-fractionated with a 500 bp cutoff,
using GibcoBRL Life Technologies' cDNA Size Fractionation
column. The column eluent was then ligated into
Stratagene's pBluescript(tm) II XR Predigested vector
(pBluescript II SK+) that has been digested with EcoRI
and XhoI, and phosphorylated by Stratagene). 97% of the
white and blue colonies appear to contain recombinant
plasmids with cDNA inserts, based on size (n=30). This
library was constructed by Dr. Paul Keim and Dr. Virginia
Coryell."

ORIGIN

Alignment Scores:
Pred. No.: 1.4e-71 Length: 500
Score: 834.00 Matches: 154
Percent Similarity: 97.59% Conservativity: 8

Best Local Similarity: 92.77% Mismatches: 4
 Query Match: 26.74% Indels: 0
 DB: 10 Gaps: 0

US-09-980-364-2 (1-579) x AW780688 (1-500)

```

QY 206 GlnArgThrSerIleTyrArgGlyValThrArgHisArgTTPThrGlyArgTyrGluAla 225
Db 3 GAGCGCACCTCCATCTACCGCGCGTCCACCCACATAGATGCGGAGGATACGAAGCT 62
QY 226 HisLeuTTPAspAsnSerCysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyr 245
Db 63 CATCTATGGACAATAGTGTGTAAGAGAGAGGCCAAAGCAGGAAAGGAGACAAGTTTAC 122
QY 246 LeuGlyGlyTyrAspLysGluGlyLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLys 265
Db 123 CTGGGTGTTATGACAAGAGATAGGAGCCAGCGGCTTACGATCTCGCAGCTCTCAAG 182
QY 266 TyrTTPGlyThrThrThrAsnPheProMetSerGluTyrGluLysGluValGlu 285
Db 183 TACTGGGTCCCAACTACCAACCACTTCCCAATTCCAACTATGAGAGGAACTGGAG 242
QY 286 GluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgArgLysSerSerGly 305
Db 243 GAGATGAAGACATGACCAGGCAAGAGTTGTGCTTCTACGAGGAGAGAGAGAGTGGT 302
QY 306 PheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHisGlyArgTyr 325
Db 303 TTCTCTAGGGGGCCCTCTATATACAGAGAGTGTGAGCAGACACCAACCATGGCCGATGG 362
QY 326 GlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThr 345
Db 363 CAGCGAGATAGGAGAGTTCGCGGAAACAAAGACCTCTACCTTGGAACTTTCAGCACC 422
QY 346 GlnGluGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAla 365
Db 423 CAAGAAGAAGCTGCTGAGCGCTATGACATTCGCTATCAAAATTCAGGGGATTAATGCA 482
QY 366 ValThrAsnPheAspMet 371
Db 483 GTCACAAACTTTGACATG 500

```

Search completed: July 19, 2004, 18:27:44
 Job time : 4169 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 19, 2004, 10:22:57 ; Search time 6533 Seconds
(without alignments)
3841.362 Million cell updates/sec

Title: US-09-980-364-2
Perfect score: 3119
Sequence: 1 MNNWLGFLSLPYEQNHHRK.....YHGGGGEVAPTFTVNDN 579

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0980364/runat_15072004_095235_25866/app.query.fasta_1.775
-DB=GenEmbl -QFMT=fastap -SUFFIX=ige -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pcst -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-CUTFM=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0980364@cgn2_1_4198@runat_15072004_095235_25866 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_hug.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pi.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

29: em_vi.*
30: em_hug_hum.*
31: em_hug_inv.*
32: em_hug_other.*
33: em_hug_mus.*
34: em_hug_pln.*
35: em_hug_rpd.*
36: em_hug_mam.*
37: em_hug_vrt.*
38: em_sy.*
39: em_hugo_hum.*
40: em_hugo_mus.*
41: em_hugo_other.*

Result No.	Score	Query Match	Length	DB ID	Description
1	3119	100.0	2014	6	AX058687 Sequence
2	3119	100.0	2014	8	AF317904 Brassica
3	3041	97.5	1992	8	AF317905 Brassica
4	3041	97.5	2011	6	AX058689 Sequence
5	2690.5	86.3	4873	6	AX058691 Sequence
6	2690.5	86.3	4873	6	AX058691 Sequence
7	2096	67.2	33563	8	AT1086 Arabidops
8	2073.5	66.5	5151	6	AX058692 Sequence
9	2073.5	66.5	5151	8	AF317907 Arabidops
10	1008	32.3	2546	8	AY461432 Nicotiana
11	994	31.9	2153	8	AK111891 Oryza sat
12	973.5	31.2	1926	6	AX555220 Sequence
13	973.5	31.2	2510	8	AK106306 Oryza sat
14	965.5	31.0	2164	8	AY062179 Oryza sat
15	963	30.9	1745	8	AK101959 Oryza sat
16	942	30.2	2010	6	AX555223 Sequence
17	942	30.2	2344	6	AX555216 Sequence
18	930	29.8	1773	8	ZMWHC1 247554 Z.mays mrna
19	924	29.8	2523	6	AX555218 Sequence
20	913.5	29.3	1738	6	BD274518 Methods f
21	913.5	29.3	1738	6	AR316369 Sequence
22	913.5	29.3	1738	6	AR427903 Sequence
23	891	28.6	1699	8	AY117207 Arabidops
24	891	28.6	1905	8	ATU41339 Arabidops
25	891	28.6	2056	8	AY080706 Arabidops
26	891	28.6	2148	6	BD274516 Methods f
27	891	28.6	2148	6	AR316367 Sequence
28	891	28.6	2148	6	AR427901 Sequence
29	891	28.6	2148	8	ATU40256 Arabidops
30	891	28.5	1955	8	ATU4028 Arabidops
31	809	25.9	1660	8	AK109839 Oryza sat
32	798	25.6	105807	8	AC006085 Arabidops
33	771	24.7	1749	6	AX048247 Sequence
34	752	24.1	81677	8	AB025629 Arabidops
35	744	23.9	153439	8	AP003313 Oryza sat
36	741.5	23.8	94788	2	AC144930 Medicago
37	727.5	23.3	1440	8	AK109848 Oryza sat
38	709.5	22.7	4024	8	AY062180 Oryza sat
39	685	22.0	1069	8	AY133880 Arabidops
40	685	22.0	1237	8	AY045915 Arabidops
41	677	21.7	103788	8	AP004747 Sequence
42	674.5	21.6	1572	8	AK111617 Oryza sat
43	670.5	21.5	110766	8	ATT28119 Arabidops
44	670.5	21.5	198493	8	ATCHRIV88 Arabidops
45	667	21.4	1866	8	AK106769 Oryza sat

ALIGNMENTS

RESULT 1

AX058687
LOCUS AX058687 2014 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 1 from Patent WO0075330.
ACCESSION AX058687
VERSION AX058687.1 GI:12311028
KEYWORDS Brassica napus (rape)
ORGANISM Brassica napus
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1
AUTHORS Boutiller, K., Ouellet, T., Custers, J., Hattori, J., Miki, B. and van
lockeren Campagne, M.
TITLE Use of the bnm3 transcriptional activator to control plant
embryogenesis and regeneration processes
JOURNAL Patent: WO 0075330-A1 14-DEC-2000;
Plant Research International (NL); THE MINISTER OF AGRICULTURE
(CA)
FEATURES
source Location/Qualifiers
1..2014
/organism="Brassica napus"
/mol_type="unassigned DNA"
/db_xref="taxon:3708"
ORIGIN
Alignment Scores:
Pred. No.: 1..61e-214 Length: 2014
Score: 3119.00 Matches: 579
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-980-364-2 (1-579) x AX058687 (1-2014)
QY 1 MetAsnAsnTrrLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLys 20
DB 120 ATGAATAATACTGTTAGGCTTTCTCTCTCTATGAAACAAATCACCATCGTAAG 179
QY 21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
DB 180 GACGCTACTCTTCACCAACACACGCTGTAGATGTCGCCGAGAGTACTGTACGAT 239
QY 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerPropheGly 60
DB 240 CCGACCGCTGCTCCGATGAGTCTTCAGCCATCCAAACATCGTTCTCTCCCTTTGGT 299
QY 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrrPaspilleAsn 80
DB 300 GTCGTCGTCGATGCTTTCACCAAGACACACAAATAGTCACTCCCGAGATTGGGACATCAAT 359
QY 81 GlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsnPhe 100
DB 360 GGTGTGTCATCAATAACATCCCAACCATGAGCAGATGACCAAGCTTGAGAAATTC 419
QY 101 LeuGlyArgThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 120
DB 420 CTTGGCCGACACACGATTTACACACCAACCAACCAACCAACCAACCAACCAACCAAC 479
QY 121 GlyCysTyrGlyGlyAspGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 140
DB 480 GGCTGTTATGAGGAGGAGAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 539
QY 141 TrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaLys 160
DB 540 TGGCTGAGAAATCAACCCCGGATTAATGTTGATATCAAGAAATGCAATGCTGCAAAA 599
QY 161 GlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAsnAspSerAsnAsn 180
DB 600 GGCCTGTCCCTCTCAATGAACTCATCTACTTCTTGTGATAACCAACACGACGACGAATAC 659
QY 181 AsnValValAlaGlnGlyLysThrIleAspAspSerValGluAlaThrProLysLysThr 200

DB 660 AACGTTGTTGCCCAAGGAGACATATTGATGATAGCGTTGAAGCTACACCGAAGAAACT 719
QY 201 IleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisArgTrpThr 220
DB 720 ATTGAGAGTTTGGACAGAGGACGCTCTATATACCGCGGTGTACAAAGGCATCGTGGACA 779
QY 221 GlyArgTyrGluAlaHisIleuTrpAspAsnSerCysLysArgGluGlyGlnThrArgLys 240
DB 780 GGAAGATATGAGGCACATTTATGGGATAATAGTTGTAAAGAGAGAGGCCAAACGCGCAA 839
QY 241 GlyArgGlnValTyrLeuGlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyrAsp 260
DB 840 GGAAGACAAAGTTTATTGGAGGTTATGACAAAGAAAGAAAGACGCTAGGCTTATGAT 899
QY 261 LeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrThrThrThrThrThrThrThr 280
DB 900 TTAGCCCGCACTCAAGTATTGGGAACCAACCACTACTACTTAACCTCCCATGAGCGAATAT 959
QY 281 GluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArg 300
DB 960 GAAAGAGGTAGNAGAGATGAAGCACATGACAAAGCAAGAGTATGTTGCCCTCACTGCGC 1019
QY 301 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHis 320
DB 1020 AGGAAAAGTAGTGGTTTCTCTCGTGTGCATCGATTTATCGTGGAGTAAACAGACATCAC 1079
QY 321 GlnHisGlyArgTrrPdinAlaArgIleGlyArgValAlaGluValAsnLysAspLeuTyrLeu 340
DB 1080 CAAACATGGAAGATGGCAAGCTAGGATAGGAGAGCTGCGCGGTAAACAAAGACCTCTACTTG 1139
QY 341 GlyThrPheGlyThrGlnGluAlaGluAlaGluAlaTyrAspIleAlaAlaIleLysPhe 360
DB 1140 GGAACCTTTTGGCACACAAGAGAGCTGCAGAGGCATACGACATTCGCGCCATCAAAATTC 1199
QY 361 ArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeu 380
DB 1200 AGAGATTAAACCGCAGTAGTAACTTCGACATGAACAGATCAACAGTAAAGCAATCCTC 1259
QY 381 GluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgPro 400
DB 1260 GAAAGCCTAGTCTTCTTATGTTAGCGCCGCAAAACGCTCTCAAGAGGCTAAACGTCG 1319
QY 401 ValProSerMetMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerGly 420
DB 1320 GTTCCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1379
QY 421 TrpGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeuLeuHisGlnHis 440
DB 1380 TGGCAAAACGCTCGGTTTCAGCATCATCAGGAGTAGATTTGAGCTTATTGCAACCAAT 1439
QY 441 GlnGluArgTyrAsnGlyTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyr 460
DB 1440 CAAGAGAGTACAATGGTTTATTATTAATGAGGAAACTTGTCTTCGAGAGTCTAGG 1499
QY 461 AlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeuMet 480
DB 1500 GCTTGTTCAAACAGAGAGATGATCAACCATTTCTTGACCAACACGACGAGCTCATG 1559
QY 481 ThrAsnIleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsnVal 500
DB 1560 ACTAATATCGATCATCAAAAGTTCTGTTTCGGATGATTCGGTTACTGTTGTGGAAATGTT 1619
QY 501 ValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAspAlaTyrAlaAla 520
DB 1620 GTTGTTTATGGTGGTATCAAGGATTTTCAGCCCGGTTAACTCGATGCTCGCTGCT 1679
QY 521 SerGluPheAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGln 540
DB 1680 AGTGAGTTTGATTAACGCAAGAAACCATTAATTTACTTTGCTCAGCAGCAGCAGCCAG 1739
QY 541 GlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyr 560


```

Db      1740 CAGTCCAGGTGGAGATTTCCTCCGGCAATGACGAATAATGTTGGCTCTAATATGAT 1799
QY      561 TyrHisGlyGluGlyGlyGluValAlaProThrPheThrValTyrAsnAspAsn 579
Db      1800 TACCATGGGAGGTGGTGGAGAGTGTCTCCACATTTTACAGTTTGGACGCAAT 1856

RESULT 2
AF317904      2014 bp      mRNA      linear      PLN 10-OCT-2002
LOCUS      Brassica napus AP2/EREBP transcription factor BABY BOOM1 (BBM1)
DEFINITION      mRNA, complete cds.
ACCESSION      AF317904
VERSION      AF317904.1
KEYWORDS      GI:21069050
SOURCE      Brassica napus (rape)
ORGANISM      Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE      1 (bases 1 to 2014)
AUTHORS      Boutillier,K., Offringa,R., Sharma,V.K., Kieft,H., van
              Lammeren,A.A.M., Ouellet,T., Zhang,L., Hattori,J., Liu,C.-M.,
              Miki,B.L.A., Custers,J.B.M. and van Lookeren Campagne,M.M.
              Ectopic expression of BABY BOOM triggers a conversion from
              vegetative to embryonic growth
              Plant Cell 14 (8), 1737-1749 (2002)
JOURNAL      Plant Cell 14 (8), 1737-1749 (2002)
MEDLINE      22163009
PUBMED      12172019
REFERENCE      2 (bases 1 to 2014)
AUTHORS      Boutillier,K., Ouellet,T. and Hattori,J.
TITLE      Direct Submission
SUBMITTED      (02-NOV-2000) Business Unit Plant Development and
              Reproduction, Plant Research International, Droeendaalsesteeg 1,
              Wageningen 6708 PB, The Netherlands
JOURNAL      Location/Qualifiers
FEATURES      1..2014
              /organism="Brassica napus"
              /mol_type="mRNA"
              /cultivar="Topas"
              /db_xref="taxon:3708"
              1..2014
              /gene="BBM1"
              120..1859
              /gene="BBM1"
              /codon_start=1
              /product="AP2/EREBP transcription factor BABY BOOM1"
              /protein_id="AA03800.1"
              /db_xref="GI:21069051"
              /translation="MNNWLGFSLSPEYQNHHRKDYSSTTTTTVDVAGEYCYDPTAA
              SDSSAIOFSPFPGVVDFAFRDNNHSRWDINGCCACNNIHNDQCPKLENFLG
              RRTTIYNTNENVGSGCYGGGGGSLGSLMKTWLRNOPVDNQNENAAK
              GLSLMNSSTSCDNNDSNNVVAQGTIDDSVEATPKTIESFGORTSIYRGVTRHP
              WGRYEAHLNDNSCKRGQTRKGRQVYLGVDKEKAAYDLAALKYNGTNTNRP
              MSYEKEVEEMKMTROEYVASLRKSGSFGSRGASIRYGRVTRHHQGRQALGRVAG
              NKDLYLGTFTQEEAAEAYDIAAKFRGLTAVTNFDNNRYNVAIKLLESPLPIGSAK
              RLKEANRPVPMIMINNVSESNASQWQNAVOHQGVDSLHLHQHQRVNGYYN
              GNLSSRSACAKFKQDQHHLSNTQSLMTNIDHQSVDSDSVTCGNVGVGGYG
              FAAPVNDVAAASEFDYNARNHYVFAQQOQTOQSPGDFPAAATNNVGSNNVYHGGG
              GEVAPTITVNDN"
ORIGIN
Alignment Scores:
Pred. No.:      1.61e-214      Length:      2014
Score:      3119.00      Matches:      579
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      8      Gaps:      0

US-09-980-364-2 (1-579) x AF317904 (1-2014)

QY      1 MetAsnAsnAnTrrPLeuGlyPheSerLeuSerProTrrGluInAsnHisHisArgLys 20

```

```

120 ATGAATAATAACTGGTTAGGCTTTCTCTCTCTCTATGAACAAAATCACCATCGTAAG 179
21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
180 GACGCTACTCTCCACCACCAACCGTGGTAGATGTCGCGGAGAGTACTGTTACGAT 239
41 ProThrAlaAlaSerAspGluSerSerAlaAlaGlnThrSerPheProSerProPheGly 60
240 CCGACCGGTGCTCCGATGAGTCTTCAGCCATCCAAACATCGTTCTCTCTCTCTTGGT 299
61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrrPaspIleAsn 80
300 GTCGTCGTCGATGTTTCACACAGAGACAACATAGTCACCTCCCGAGATTGGGACATCAAT 359
81 GlyCysAlaCysAsnAsnLleHisAsnAspGluInAspGlyProLysLeuGluAsnPhe 100
360 GGTTCGTCATCAATAACATCCACACGATGAGCAAGATGACCAACAGCTTGAGAAATTC 419
101 LeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGlySer 120
420 CTTGGCCGCCACACACAGATTACACACCAACCAACGCTGGAGATGGAAGTGAAGT 479
121 GlyCysTyrGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMetIleLysThr 140
480 GGTGTTATGAGAGAGAGACGGTGGTGGCTCACTAGGACTTTTCGATGATAAGACA 539
141 TrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaLys 160
540 TGGCTGAGAAATCAACCCGTCGATAATGTTGATTAATCAAGAAATGGCAATGCTGCAAA 599
161 GlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAsnAspSerAsnAsn 180
600 GGCCTGTCCCTCAATCAACTCATCTACTTCTTGTGTATAACAACAACAGACGCAATAAC 659
181 AsnValValAlaGlnGlyThrIleAspAspSerValGluAlaThrProLysLysThr 200
660 AACGTTGTTGCCCAAGGAGACATTTGATGATGCGTTGAAGCTACACCGAAGAAACT 719
201 IleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisArgTrrThr 220
720 ATTGAGAGTTTGGACAGAGACGCTATATACCGCGGTGTACAAAGCATCGTGGACA 779
221 GlyArgTyrGluAlaHisLeuTrrPaspAsnSerCysLysArgGluGlyGlnThrArgLys 240
780 GGAAGATATAGGCACATTTATGGGATAATAGTTGTAAGAGAGAGGCGCAACCGCGCAA 839
241 GlyArgGlnValTyrLeuGlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTrrAsp 260
840 GGAAGACAGATTATTTGGAGGTTTACAAAGAGAAAGACGCTAGGCTTAGGCTTATCAT 899
261 LeuAlaAlaLeuLysTyrTrrPglyThrThrThrThrAsnPheProMetSerGluTyr 280
900 TTAGCGGCACCTCAAGTATTGGGAACCACTACTACTACTAACTTCCCATGAGCGAATAT 959
281 GluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArg 300
960 GAAAAGAGGTAGAAGATGAAGACATGACAAAGAGAGATATGTTGCTCTCACTGGCG 1019
301 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHis 320
1020 AGAAAAGTAGTGGTTCTCTCGTGGTCATCGATTTATCGTGGAGTAAACAAGACATCAC 1079
321 GlnHisGlyArgTrrPglAlaAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTrrLeu 340
1080 CAACATGGAAGATGGCAGCTAGGATAGGAAGAGTCGCGGTAACAAGACCTCTACTTGG 1139
341 GlyThrPheGlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaLysPhe 360
1140 GGAACCTTTTGGCACACAGAGAGAGCTGCAGAGGATACGACATTCGCGGCATCAATTC 1199
361 ArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaLysLeu 380
1200 AGAGGATTAAACCGCATGACTAACTTCGACATGAACAGATACAACGTTTAAAGCAATCTTC 1259

```


Db 651 ATTGAGAGTTTGGACAGAGGAGCTCTATATACCGGGTGTACAGGCATCGGTGGACA 710
QY 221 GlyArgTyrGluAlaHisLeuTrpAsnSerCysLysArgGluGlnThrArgLys 240
Db 711 CGAAGATATGAGCACATTTATGGGATAATAGTTGTAAACGAGAGGCGCAACGCGCAAA 770
QY 241 GlyArgGlnValTyrLeuGlyGlyTyrAspLysGluLysAlaAlaArgAlaTyrAsp 260
Db 771 GGAAGCAAGTTTATTTGGAGGTTTNGCAAGAGAAAAGACAGCTAGGCGCTTATGAT 830
QY 261 LeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrAsnPheProMetSerGluTyr 280
Db 831 TTAGCCGCACTCAAGTATTTGGGAACCAACCACTACTACTAACTTCCCATGAGCGAATAT 890
QY 301 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHis 320
Db 951 AGCAAAAGTAGTGTCTCTCGTGGTGCATCGATTATCGTGGAGTAAACAAGACATCAC 1010
QY 321 GlnHisGlyArgTrpGlnAlaAlaGlyArgValAlaGlyAsnLysAspLeuTyrLeu 340
Db 1011 CAACATGGAAGATGGCAAGCTAGGATAGGAAGAGTCCCGGTAACAAGACCTCTACTTG 1070
QY 341 GlyThrPheGlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPhe 360
Db 1071 GGAACCTTTGGCACACAGAGAGAGCTGCAGAGGCATACGACATTCGGGCCATCAATTC 1130
QY 361 ArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeu 380
Db 1131 AGAGGATTAAACCCAGTGACTAACTTCGACATGACAGATACAACTGTTAAAGCAATCCCTC 1190
QY 381 GluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgPro 400
Db 1191 GAAAGCCCTAGTCTTCTATTTGGTGGCCGCAACAGCTCTCAAGAGAGCTAACCGTCCG 1250
QY 401 ValProSerMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerGly 420
Db 1251 GTTCCAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1310
QY 421 TrpGlnAsnAlaAlaValGlnHisHisGlnGlyValAspLeuSerLeuLeuHisGlnHis 440
Db 1311 TGGCAAAACGCTCGGTTTCAGCATCATCAGGAGTAGATTGAGCTTATTTGCGAGCAACAT 1370
QY 441 GlnGluArgTyrAsnGlyTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaArg 460
Db 1371 CAAGAGAGGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1430
QY 461 AlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeuMet 480
Db 1431 GCATTGTTTCAACAAGAGGATGATCAACACCATTTCTTGTAGCAACACGACGACGCTCATG 1490
QY 481 ThrAsnIleAspHisGlnSerSerValSerAspAspSerValThrValCysGlyAsnVal 500
Db 1491 ACTAATATCGATCATCAAGATTCTGTTTTCAGATGATTCGGTTACTGTTTGTGGAATGTT 1550
QY 501 ValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAspAlaTyrAlaAla 520
Db 1551 GTTGTTATGTTGTTATCAAGGATTTCAGAGCCCGGTTAACTGCGATGCTACGCTGCT 1610
QY 521 SerGluPheAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGln 540
Db 1611 AGTAGTTTGTACTATAACGCAAGAAACCATTTATTTACTTTTGTCTAGCAGCAGCAGCCAG 1670
QY 541 GlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyr 560
Db 1671 CATTCGCCAGGAGGAGATTTTCCCGGCAATGACGAATAATGTTGGCTCTAATATGAT 1730
QY 561 TyrHisGlyGluGlyGlyGlyGluValAlaProThrPheThrValTrpAsnAspAsn 579
Db 1731 TACCATGGGGAAGTGGTGGAGAAGTTGCTCCACATTTTACAGTTTGAACGACCAAT 1787

RESULT 4
AX058689
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
Alignment Scores
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
US-09-980-364-2 (1-579) x AX058689 (1-2011)
QY 1 MetAsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLys 20
Db 51 ATGAATAATAAATCTGGTTAGGCTTTCTCTCTCTTATGAACAAAATCACCATCGTAAG 110
QY 21 AspValTyrSerSerThrThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
Db 111 GACGTCGCTCTTCCACCAACCAACCCCGCTAGATGTCGCGGAGAGTACTGTTACGAT 170
QY 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
Db 171 CCGACCGCTGCTCGGATGAGTCTTCAGCCATCCAAACATCGTTTCTTCTCCCTTTGGT 230
QY 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAspIleAsn 80
Db 231 GTCGTCCTCGATGCTTTTCCACGAGACAAACAATAGTCACTCCGAGATGGGACATCAAT 290
QY 81 GlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsnPhe 100
Db 291 GGTAGTGATGATGAGGAGAGCGTGGTGGTCTCACTAGGACTTCGATGATAAAGACA 470
QY 101 LeuGlyArgThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 120
Db 351 CTTGGCCGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 410
QY 121 GlyCysTyrGlyGlyAspGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 140
Db 411 GGGTGTATGAGGAGGAGAGCGTGGTGGTCTCACTAGGACTTCGATGATAAAGACA 470
QY 141 TrpLeuArgAsnGlnProValAspAsnValAspAsnValAspAsnValAspAsnValAsp 160
Db 471 TGGCTGAGAAATCAACCCCGTGGATATGTTGATATCAAGAAATGGCAATCGTGCATAA 530
QY 161 GlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSerAsnAsn 180
Db 531 GGCTGTCTCTCTCAATGAACTCACTACTTCTTCTGTATTAACAACAACACTACAGCATAC 590

AX058689 Sequence 3 from Patent WO0075330.
AX058689 AX058689.1 GI:12311029
Brassica napus (rape)
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 Boutlier,K., Ouellet,T., Custers,J., Hattori,J., Miki,B. and van Lookeren Campagne,M.
Use of the bnm3 transcriptional activator to control plant embryogenesis and regeneration processes
Plant: WO 0075330-A 3 14-DEC-2000;
Patent Research International (NL) ; THE MINISTER OF AGRICULTURE (CA)
Location/Qualifiers
1. .2011
/organism="Brassica napus"
/mol_type="unassigned DNA"
/db_xref="taxon:3708"

Db 1920 ACAAAGACCAATACACACGACGATACCTAGCCCTATTTTTCCTGGGCTTATTATCG 1979
QY 75 ----- 75
Db 1980 ATTTCATTATTTTGAGAAATACATGTGTGGGTTTGATGTGTTGTCATATAGTAAT 2039
QY 75 ----- 75
Db 2040 ACTAAACATATGCCAGTTATACATAGATTTTAAAAAGATATACATGGATGAAATG 2099
QY 75 ----- 75
Db 2100 AAATTTGACATTTCCCTTTATTCATATATCATATATCATCATCATGTGTACCTTT 2159
QY 75 ----- 75
Db 2160 TGATTTGTATATTTGTTCTTACAGTTGAAGGAGAGATAACCAATATACCCATTTGTATA 2219
QY 75 ----- 75
Db 2220 TTATAGATCGGTGATGAAAAAGTAAATTTAAACAAATTATGATATATAGGCCATTATCTT 2279
QY 76 -----
Db 2280 TGATTTTCTTTTATAGATTGGGACATCAATGGTTGTGATGCAATTAACATCCACAAC 2339
QY 90 AspGluGlnAspGlyProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsn 109
Db 2340 GATGAGCAAGATGCCAAAGCTTGAGAAATTCCTTGGCGCACCAACCCAGCATTTTACAAC 2399
QY 110 ThrAsnGluAsnValGlyAspGlySerGlySerGlyCysTyrGlyGlyGlyGly 129
Db 2400 ACCAACGAAACGTTGGAGATGGAAGTGGAGTGGCTGTTATCGAGGAGGAGCGGTGT 2459
QY 130 GlyGlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProValAspAsn 149
Db 2460 GGTGGCTCACTAGACATTCGATGATAAGACATGCTGAGAAATCAACCCGTGGATAAT 2519
QY 150 ValAspAsnGlnGluAsnGlyAsnAlaLysGlyLeuSerLeuSerMetAsnSer 169
Db 2520 GTTGATAATCAAGAAATGCAATGCTGCAAAAGGCGCTGCTCTCAATGAACTCATCT 2579
QY 170 ThrSerCysAspAsnAsnAspSerAsnAsnValValAlaGlnGlyLysThrIle 189
Db 2580 ACITCTGTGATAACAACAGCAGCAATAAACAGCTGTGTCCTCAATGAACTCATCT 2639
QY 190 AspAspSerValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSer 209
Db 2640 GATGATAGCGTTGAAGCTACACCGAAGAAACTATTGAGAGTTTGGACAGGACGTCT 2699
QY 210 IleTyrArgGlyValThr----- 215
Db 2700 ATATACCGCGGTGTAC- AAGGTGCCCTTCATTATTAAATGCTAAAGTCTCG 2758
QY 215 ----- 215
Db 2759 CTTGAATTTGTTATCTTCTGTGTAAGTCTGGACATTTGATGCTCTGTGGGAGA 2818
QY 215 ----- 215
Db 2819 GTGCTACCGAATGCTCTGTATATATAGTATCAAGAGAGATATTGTTATTTGGGCTTA 2878
QY 215 ----- 215
Db 2879 TATAGATAATACATATATATATATATATACATGCTAGCTGTTGATGACATGTATGTCG 2938
QY 216 -----
Db 2939 TATTAATATGATAGGCATCGGTGGACAGGAAGATATGAGGCACATTTATGGGATAATAGT 2998
QY 232 CysLysArgGluGlyGlnThrArgLysGlyArgGln----- 243
Db 2999 TGTAAGAGAGAGGCCAAACCGCAAGGAGAGACAAGGTATATATATATATATATATATAT 3058

QY 243 ----- 243
Db 3059 TTGATCATATTTTCATACAGATTTACTTTCAAACATAATATAGCTTTTTCGATCATTTGT 3118
QY 243 ----- 243
Db 3119 CATGTTTTTATCAAAATTTGCACCTGGTGTCTCTCAGTTTATTTGGGTAAGTAAT 3178
QY 243 ----- 243
Db 3179 TTATTATAAATTTGGACGAAGCTGTGATGGTAAATCTAAATATATAATCAAAATTTGTTT 3238
QY 244 -----ValTyrLeu----- 246
Db 3239 ATTTTGTGTATACATTTATATATAATCAAAATAGCGATACGATCTACATTCATTTGT 3298
QY 247 -----GlyGlyTyrAspLysGluGlyLysAlaLysAlaLysAspLeu 261
Db 3299 TGTCTATATCATCGAGGAGTTATGACAAAGAAAAAGCAGCTAGGCTTATGATTTA 3358
QY 262 AlaAlaLeuLysTyrTyrGlyThrThrThrThrAsnPhePro----- 276
Db 3359 GCCGCACTCAAGTATTTGGGAAACCACTACTACTAACTTCCCGTAAGTCAATCAATG 3418
QY 276 ----- 276
Db 3419 TTGTACAAGATTTCAATACTTAGAACCAATTTTCTTTTATTAAGATGCTATTATC 3478
QY 276 ----- 276
Db 3479 TTATTATTAAATGGCATGTTTATATCGTTACATTTATTACAATAAAAAAGTACTTTTGGT 3538
QY 277 -----MetSerGluTyrGluLysGluValGluMetLysHisMetThr 291
Db 3539 TGATATAATATGATAGTGGCAATATGAAAAGAGGTAGAGAGATGAAGCACATGACA 3598
QY 292 ArgGlnGluTyrValAlaSerLeuArgArg----- 301
Db 3599 AGGCAAGATATGTTGCTCTCACTGCGCAGGTATATAATGAACCTTCTGTATATTATTGCAT 3658
QY 301 ----- 301
Db 3659 ATGGCATCTTATTATATACATGTATATTAGTATTATTATATAGAACCCATCACCTCAC 3718
QY 301 ----- 301
Db 3719 GTTTATATTTAAANATGTCCGTATTCCAGTCAGATTATCAGCATACACCTATATATAA 3778
QY 302 -----LysSerSerGlyPheSerArgGlyValaSerIleTyrArg 314
Db 3779 TAGACATTTAAATATATGCAGGAAAGT-AGTGGTTTCTCTCGTGGTGCATCGATTATCGT 3837
QY 315 GlyValThr----- 317
Db 3838 GGAGTAAC- AAGGTATTATACAGAGAGAACGAATCCTATTTTGTAGTACATATATAT 3896
QY 317 ----- 317
Db 3897 ATAAAAATATAATATAAGATATCACATTTTATATATTAATGAATATTTCTTAAATGGGTCC 3956
QY 318 -----ArgHisHisGlnHisGlyArgTrpGlnAlaArgLysValaLysValaLysAsnLys 336
Db 3957 AAAAGACATCAACCAATGGAAGATGGCAAGCTAGTAGAGAGAGTGCCTGGTAACAA 4016
QY 337 AspLeuTyrLeuGlyThrPhe----- 343
Db 4017 GACCTTCTACTTGGGAAC-TTTTGGTACGTTTAGTCTTCTTACTTAAACTTCACAATCAA 4075
QY 343 ----- 343
Db 4076 ATCTATAACAAAGATATCAACTAAACAACTACAATATATCTAAGTAAGTGTACATAT 4135

QY 75 ----- 75
Db 28880 CATTTTCGTATATTATAGATCGGTGATGAATCAAAACCCCAACCGTTGGGGATCGCTTTGTTTT 28821
QY 76 ----- AspTrpAspIleasnGlyCysAlaCysAsnAsnIleHisAsnAspGluGln 92
Db 28820 TTGCTCTATAGATTGGGACATCAATGGTGGTCATGCAATAACATTAACAATAACGAACAA 28761
QY 93 AspGlyProLysLeuGluAsnPheLeuGlyArgThrThrIleTyAsnThrAsnGlu 112
Db 28760 AATGGACCAAAAGCTTGAGAAATTTCTCGCGCCGACCCACCGATTTACAAATACCAACGAG 28701
QY 113 AsnValGlyAspGlySerGlySerGlyCysTyArgGlyGlyAspGlyGlyGlySer 132
Db 28700 ACCGTTGTAGATGGAATGGCGAT---TGCGAGGAGGAGACGGTGGTGGCGCGCTCA 28644
QY 133 LeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProValAspAsnValAspAsn 152
Db 28643 CTAGGCCCTTCGATGATTAACAACATGGCTGAGTAATCATTCGTTGCTAATGCTAATCAT 28584
QY 153 GlnGluAsnGlyAsnAlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCys 172
Db 28583 CAAGCAATGGTAAACGGTGCAGAGGCTTGTCCCTCTCTATCAATTCATCTACTAGT--- 28527
QY 173 AspAsnAsnAspSerAsnAsnValValAlaGlnGlyLysThrIleAspAspSer 192
Db 28526 GATAGCAACACTACACAAACATGATGATGCTGCTCCAGAGAGACTATTGTTGATGTC 28467
QY 193 ValGluAlaThrProLysThrIleGluSerPheGlyGlnA-gThrSerIleTyArg 212
Db 28466 GTAGAAACTACACCGAAGAAACTATTGAGAGTTTGGACAAAGCAAGCTCTATATACCGC 28407
QY 213 GlyValThr----- 215
Db 28406 GGTGTTAC-AAGGTTAATTTCAATCATCTATGTATATTTTATTGCTGCTAAATTTGAT 28348
QY 215 ----- 215
Db 28347 TTTCCTGGTATTGTTGGACATTTCTAATGTTGGTTGAGAGAGAGTGCAACGGAATGT 28288
QY 215 ----- 215
Db 28287 CTCTCAATGTATATTAAGAGAAACATTAATAGTGTACATGGGTTTATATATACAATAA 28228
QY 216 ----- Arg 216
Db 28227 TACGTCATATATATGTTGCTCTTGATCATAGTATATATATGTTGAATTTAATGTCAGG 28168
QY 217 HisArgTrpThrGlyArgTyGluAlaHisLeuTrpAspAsnSerCysLysArgGlyGly 236
Db 28167 CATCGTGCAGAGGTAGATACAGGCAATTTATGGGACATAGTTGCCAAAAGAGAAGGC 28108
QY 237 GlnThrArgLysGlyArgGlnValTy----- 245
Db 28107 CAGACTCCCAAGGAGACAGG-TACTATATATAAAGCTAATTTTAAATTTTCATT 28049
QY 245 ----- 245
Db 28048 TACCATTATTTTCAAACTAATTTAGGTTTCTTTTCATGTGTTTCATCAAAATTTGCAC 27989
QY 245 ----- 245
Db 27988 CTGATGGCTCTCTTTTCAAGTTTATCTCGGTAAGTTCTTGATTTTAAGCTATATAATTAATA 27929
QY 246 ----- Leu 246
Db 27928 ATAGATGACTATTAAATCTATTCTAAGCAAAATATAATTTGTTGTTATCTGATCTTACA 27869
QY 247 GlyGlyTyAspLysGluGlyAlaAlaArgAlaTyAspLeuAlaAlaLeuLysTy 266
Db 27868 GGAGTTTATGAAGAAGAAAGAGAGCTAGGCTTACGATTTAGCCGCACATAAGTAT 27809

QY 267 TrpGlyThrThrThrThrThrAsnPhe-Pro----- 276
Db 27808 TGGGGAACCAACCACTACTACTAACTTCCCGTATGTTAATTAATCAATAATATATACATA 27749
QY 276 ----- 276
Db 27748 AATTCCTAACTTCTAACCAATTTTAGTCTGAATAATGCCAATCTCTTAACTAGTATTAT 27689
QY 276 ----- 276
Db 27688 CTTACTATTAACTGTCATGTTTATATTGTTACATTTGTTACAAATAAAATAGTAATGTT 27629
QY 277 ----- MetSerGluTyGluLysGluValGluGluMetLysHisGly 290
Db 27628 GGTGGATATAATATTTCAGTTGAGTGAATATAGAAAGAGGTAGAGAGATGAACACAT 27569
QY 290 tThrArgGlnGluTyValAlaSerLeuArgArg----- 301
Db 27568 GACGAGGCAAGAGTATGTTGCTCTCTCGCAGGTACAGAAATGAAACTCTTTGAATTTATT 27509
QY 301 ----- 301
Db 27508 GCATTTTAGAAACCCCATCAGCTATATATTTTATAAATATATCGTAACATTGAATAATC 27449
QY 302 ----- LysSerSerGlyPheSerArg 308
Db 27448 ATTATTGGAAAGATATAAGAACATGTAATATATCGAGGAAAGT-AGTGGTTTCTCTCG 27390
QY 308 gGlyAlaSerIleTyArgGlyValThr----- 317
Db 27389 TGGTGCATCGATTATTCGAGGAGTAACAAGGTACGTATAATCCATCTAGATATGGAACGA 27330
QY 317 ----- 317
Db 27329 ATACAGTGTTCATTATTTTTCATGTATACATAATAATTTGTCATACATACTATTAA 27270
QY 318 ----- ArgHisHisGlnHisGlyArgT 325
Db 27269 TCTAATCTAATTAATATTTCCTTTTAAATGTTTCCAAAAGGCATCACCACATGGAAGT 27210
QY 325 rGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyLeuGlyThrPhe----- 343
Db 27209 GCAAGCTAGGATCGAAGATCGCGGTAAACAAAGACTCTCTACTTGGAACTTTTCGTA 27150
QY 343 ----- 343
Db 27149 CATTTTCCAATAAATCTATATCTATATAAGATATTAAATATATACAAATATATCTAAGTG 27090
QY 344 ----- GlyThrGlnGluAlaAlaGluAlaTyAspIleAla 357
Db 27089 AATCATACAAATATTGTAGGCACACAGGAAGGCTGCTGAGGCTTATGACATTCGACGCC 27030
QY 358 IleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyAsnValLys 377
Db 27029 ATTAAATTCAGAGGATTAAGCCAGTGAATCTTCGACATGAACAGATACATGTTAA 26970
QY 378 AlaIleLeuGluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAla 397
Db 26969 GCAATCTCGAGAGCCGAGTCTACCTATTGTTAGTTCTGCGAAACGCTCTCAAGGACGTT 26910
QY 398 AsnArgProValProSerMetMetIleSerAsnAsnValSerGluSerGluAsnSer 417
Db 26909 AATAATCCGTTTCCAGCT---ATGATGATTAGTAATAACGTTTCAGAGAGTGCATAAT 26853
QY 418 AlaSerGlyTrpGlnAsnAlaAlaValGlnHisHisGlnGlyValAspLeuSerLeu 437
Db 26852 GTTAGCGGTGCAAAACACTGCGTTTCAGCATCATCAGGAATGAGTTTTCAGCTTATTG 26793
QY 438 HisGlnHisGlnGluArgTyAsnGlyTyTrpTyAsnGlyGlyAsnLeuSerSerGlu 457
Db 26792 CAGCAACAGCAGGAGGTACGTTGGT---TATTACATATGAGGNAACCTTCTTACCAG 26736
QY 458 SerAlaAla-gAlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGln 477

Db 3042 TTCTTGGTATGTTTGGGACATCTTAATGGTTCCGTTGAGAGAGAGTGCAACGGAATGT 3101
QY 215 ----- 215
Db 3102 CTCTCAATGATATTAAGAGAGAAACATTAATAGTGATACATGGGTTTATATATACAATAA 3161
QY 216 -----Arg 216
Db 3162 TACGTCATATATATCGGTATGCTCTTGATCATAGTATATATATGTTGAATTTAATGTGAGG 3221
QY 217 HisArgTprThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGluGly 236
Db 3222 CATCGGTGACAGGTAGATACGAGGCACATTTATGGGACAATAGTTGCAAAAGAGAAGGC 3281
QY 237 GlnThrArgLysGlyArgGlnValTyr----- 245
Db 3282 CAGACTCGCAAGGAAGACAAGG-TACTATATATATAAAGCTAAATTTTAAATTTTCATT 3340
QY 245 ----- 245
Db 3341 TACCATTTATTTCAAACCTAATTTAGGTTTCTTTTCATGTGTTTCATCAAAATTTGCAC 3400
QY 245 ----- 245
Db 3401 CTGATGGCTCTCTTTTCAGTTTATCTGGGTAAGTTCTTTGATTTAAGCTATAAAATTAATA 3460
QY 246 -----Leu 246
Db 3461 ATAGATGACTATTAAATCTATTTCAAGCAAAATATATGTTGTGTTATCTGATCCTTACA 3520
QY 247 GlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaLeuLysTyr 266
Db 3521 GGAGGTTATGACAAAGAGAAAGACAGCTAGGCTTACGATTTAGCGGCATAAAGTAT 3580
QY 267 TrpGlyThrThrThrThrAsnPhe-Pro----- 276
Db 3581 TGGGACCCACCACTACTACTAACTTCCCGTATGTTAAATCAATAATATATACATA 3640
QY 276 ----- 276
Db 3641 AATTCCTAACTCTTAACCAATTTTAGTCTGAATAATGCCAATCTCTTAAACTAGTATTAT 3700
QY 276 ----- 276
Db 3701 CTTACTATTAACTGTCATGTTTATATGTTTACATAAAATAGTAATGTTGGTGATA 3760
QY 277 -----MetSerGluTyrGluLysGluValGluMetLysHisMetThrArgG1 293
Db 3761 TAATATTTCAGTTGAGTGAATATGAGAAAGAGGTAGAGAGATGAACACATGACGAGGCA 3820
QY 293 nGluTyrValAlaSerLeuArg----- 301
Db 3821 AGAGTATGTTGCCTCTCTGCGCAGGTACAGAACTCTTGAATTTATGCAATTTTAG 3880
QY 301 ----- 301
Db 3881 AAACCCATCACGTATATATTTTAAATATATATCGTAACATGTAATGAATCAATTTATTTGG 3940
QY 302 -----LysSerSerGlyPheSerArgGlyAlaLase 311
Db 3941 AAAGATATAAGAACATGTAATATATGACGAGAAAAGT-AGTGGTTTCTCTCGTGTGTCATC 3999
QY 311 rIleTyrArgGlyValThr----- 317
Db 4000 GATTTATCGAGAGTAAAC-AAGGTACGTATATATCCATCTAGATATGAAACGAATACTAGT 4058
QY 317 ----- 317
Db 4059 GTTTCATTATTTTGTGATGATACATAATAATATGTCATACATACTACTATTAATCTAATC 4118
QY 318 -----ArgHisGlnHisGlyArgTprGlnAl 327
Db 4119 TAATTAATATTTCTTTAAATGGTTCCAAAAGGCATCCCAACATGGAAGGTGGGGAAGC 4178

QY 327 aArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPhe----- 343
Db 4179 TAGGATCGGAAGAGTGGCGGTAAACAAAGACCTCTACTTGGGAACCTTTTCGGGTACATTTTC 4238
QY 343 ----- 343
Db 4239 CAATAAAATCTATATACATATAAGTATTAATATATACAAATATATATCAAGTGAATCATA 4298
QY 344 -----GlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaLysP 360
Db 4299 CAAATTTATGTAGGCACACAGGAAGAGGCTGCTGAGGCTTATGACATTCAGCATTTAAAT 4358
QY 360 heArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleL 380
Db 4359 TAGAGGATTAGCGCAGTACTTAATTCGACATGACAGATACATATGTTAAAGCAATCC 4418
QY 380 euGluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgP 400
Db 4419 TCGAGAGCCGAGTCTACCTATTTGCTAGTTCTCGAAACGCTCTCAAGGACGTTTAAACAATC 4478
QY 400 roValProSerMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerG 420
Db 4479 CGGTTCCAGCT---ATGATGATTAGTAAATACGTTTCAGAGAGTCCAAATTAATGTTAGCG 4535
QY 420 lyTrpGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeuLeuHisGlnH 440
Db 4536 GTTGGCAAAACACTCGGTTTCAGCATCATCAGGGAATGGATTGAGCTTATTTCAGCAAC 4595
QY 440 isGlnGluArgTyrAsnGlyTyrTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaA 460
Db 4596 ACACAGAGAGTACGCTGGT---TATTACAATGGAGGAACCTTGTCTACCGAGAGTACTA 4652
QY 460 rGAlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeuM 480
Db 4653 GGGTTTGTTCACAAAGAGGAGGAACAACACACTCTCTTGAGAAACCTCGCGAGTCA 4712
QY 480 eThrAsnIleAspHisGlnSerValSerAspSerValThrValCysGlyAsnV 500
Db 4713 TGACTAATGTTGATCATCATAGCTCGACCTCTGATGATCTGTTACCGTTTGTGGAAATG 4772
QY 500 alValGlyTyrGlyTyrGlnGlyPheAlaAlaPro-----ValAsnCysA 516
Db 4773 TTGTTAGTTATGCTGTTATCAAGATTCGCAATCCCTGTTGGAACATCGGTTAATTACG 4832
QY 516 spAlaTyrAlaAlaSerGluPheAspTyrAsnAlaArgAsnHisTyrTyrPheAla---- 534
Db 4833 ATCCCTTTACTGCTGCTGAGATTGCTTACACGCAAGAAATCATTATTACTATGCTCAGC 4892
QY 535 -----GlnGlnGlnThrGlnSerProGlyGlyAspPheProAlaAlaMetThrA 553
Db 4893 ATCAGCAACACAGCAGATTCCAGCTGCCGGGAGGAGATTTTCGGTGGCGATTTCGA 4952
QY 553 snAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGlyGluValAlaProThrP 573
Db 4953 ATAACCATAGCTCTAAATGATCTTTCACGGGAAGGTGGTGGAGAGGGGCTCCACGCT 5012
QY 573 heThrValTrpAsnAsp 578
Db 5013 TTTTCAGTTTGGAAACGAC 5029

RESULT 9

AF317907 5151 bp DNA linear PLN 10-OCT-2002
LOCUS Arabidopsis thaliana BABY BOOM (BBM) gene, complete cds.
DEFINITION AF317907
ACCESSION AF317907.1 GI:21069056
VERSION
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsais.


```
/codon_start=1
/product="ANT-like protein"
/protein_id="AAR22388.1"
/db_xref="GI:38492172"
/translation="MKKSWNDNNKSSHRSSNNNNNNNSHNSNDASNNWLFSLSP
HKWEVTSSEPOHQHCFNRSQSSALAQSYLSSSPNNVTSNSALCYGVGENNPPGHS
DSLVMPKSDGLCINEALRSHADAVOSSPKLEDFGGASQVGNHERAMALSL
DSLYHONQGYSPMCHGMYESLDETPTQISHCDAQMTAVSNGELKSGWYVAD
QHINETSVMVAGGGGGTGGGVTGNDLQSLSLSNMPSQSSCVTPRQISFNEI
ECVAITKRAGKVQKQPVHRKSIDTFQRTSOVTRHRTGRVYEAHLWDSCK
KEGCTKGRQVLYGGVDEEKAARAYDLAALKYWPSTHINFLPNYOKELMDKQMT
ROYVAHLRKSGSFGSRGSMYRVTRHHQHGWRQARIGRVAGNKLVLSTPESDEA
ARAYDVAALKFRGNVATNFIDISRDVEXIKMASNTLPAGELARRKKEPPIYNLSDH
KNEEACVQNNNNNNNDWQVSYQSSNFSNLGRNPTFSMALQDLIGLIDLNSS
NHILDDQNKYKGNHFSNASLVTSLGSSREASPSKSAASLVFAKPTNVNACIPSAQL
RPIPVGISHLPVFAALNDA"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 6,02e-63 Length: 2546
Score: 1008.00 Matches: 258
Percent Similarity: 50.08% Conservative: 65
Best Local Similarity: 40.00% Mismatches: 132
Query Match: 32.32% Indels: 190
DB: 8 Gaps: 24
```

US-09-980-364-2 (1-579) x AY461432 (1-2546)

```
Qy 2 AsnAsnThrLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLysAsp 21
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 442 AGTAGTAACCTGGTAGGATTTCTCTCACCC-----CACATGAAATGGAG 489
   |||
Qy 22 ValTyrSerSerThrThrThrValValAspValAlaGlyLeuTyrCysTyrAspPro 41
   |||
Db 490 GTTACTTCTCT-----GAACTCAACACCAACAT 519
   |||
Qy 42 ThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPhe-----ProSerPro 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 520 CATCAGTTCAATAATCAGAGCTCTGCTCTTGGCTCAAGTTTCTATCTTCTCTCTCT 579
   |||
Qy 59 PheGlyValValAspAlaPhe-----ThrArgAspAsnAsn----- 71
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 580 ATGAATGTGACTTCAATACTCTGCTCTGTGTATGTGTTGGAGAGATAATCCCTTT 639
   |||
Qy 72 SerHisser-----ArgAspTyrAspIleAsnGlyCysAlaCys----- 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 640 GGCAATCTCTTCTATCTGTATGCTCTCAAACTCTGATGCTCTCTTTCATATGGAA 699
   |||
Qy 85 -----AsnAsnLeuHisAsnAsp-----GluGlnAspGlyProLysLeuGlu 98
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 700 GCTCTCTCAAGATCTCATGAGATGCTATGCTGCAAAAGTTCTTCTCACCTAAGCTTGAG 759
   |||
Qy 99 AsnPheLeuGlyArgThr----- 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 760 GACTTTTAGTGGTGGCAATCATATGGAATCATGAAAGGAAGCTATGGCTCTTAAGC 819
   |||
Qy 105 -----ThrThrIleTyrAsnThrAsnGluAsn----- 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 820 TTAGACAGCTTATATTATCACCAAAATGAAGCTATTATTCTCCATGCATTGTCATGCC 879
   |||
Qy 113 ----- 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 880 ATGTATCAAGAAATCATTATTAGATGAACAAACCAAACTCAAAATTTCAATTGTGATGCC 939
   |||
Qy 113 ----- 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 940 CAAATGACTGCAGTGTCTGGAATGAACCTGAAAGTTGGGGCATTTATGCTATTGACCAA 999
   |||
Qy 114 -----ValGlyAspGlySerGlySerGlyCysTyrGly 124
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1000 CATATAAATGAACCTTGTAGTAGTATGTTGGTGCAGGTGGTGTGGAGTACTTCTGCT 1059
   |||
Qy 125 GlyGlyAspGlyGlyGlyGlySerLeuGlyLeuSerMetIleLysThrTyrLeuArgAsn 144
   |||
```

```
Db 1060 -----TCTGTGTGTGTACAGTTGGT----- 1080
   |||
Qy 145 GlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaLysGlyLeuSerLeu 164
   |||
Db 1081 -----GGTAATGATTTACAGTCTTTGAGCTTG 1107
   |||
Qy 165 SerMetAsn-----SerSerThrSerCysAspAsnAsnAsnAspSerAsnAsnVal 182
   |||
Db 1108 TCTATGAATCTGTGTTCTCAGCTAGCTGTGTTACTCCAAGACAAATTTCTCTCAATAGAG 1167
   |||
Qy 183 Val-----AlaGlnGlyLysThrIleAspAspSer 192
   |||
Db 1168 CTGGAATGTGTGCATTTGAAACAAAGAAAGGGCTTCTGAAAGATGTTTCAAAA--- 1224
   |||
Qy 193 ValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArg 212
   |||
Db 1225 ---CAACCTGTTTCATAGGAAATCTATTATCATACATTTGTCAGAGAACTTCTCAGTATAGA 1281
   |||
Qy 213 GlyValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTyrAspAsnSerCys 232
   |||
Db 1282 GGTGTCTACTAGGATAGGTGACAGGTAGGTATGAGCGCATTTTGTGGGATTAATAGTTGC 1341
   |||
Qy 233 LysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGlu 252
   |||
Db 1342 AAGAAGGAAGGGCAGACTAGAAAAGGAGGCAAGTTTATCTTGGGGCTATGATATGGA 1401
   |||
Qy 253 GluLysAlaAlaArgAlaTyrAspLeuAlaLysLysTyrTyrGlyThrThrThr 272
   |||
Db 1402 GAGAAAGCTGCAAGAGCATATGATTAGTCACATAAGTATTGGGACCTTCAACTCAC 1461
   |||
Qy 273 ThrAsnPheProMetSerGluTyrGluLysGluValGluGluMetLysHisMetThrArg 292
   |||
Db 1462 ATTAACCTTTCGTTAGAGAACTACCAGAAAGAGCTTGAGGATATGAAGATATGACCCGG 1521
   |||
Qy 293 GlnGluTyrValAlaSerLeuArgLysSerSerGlyPheSerArgGlyAlaSerIle 312
   |||
Db 1522 CAAGAATATGTTGCACACTTAAGAAGGAAAGTAGTGGAATCTCAAGAGGTGCTTCAATG 1581
   |||
Qy 313 TyrArgGlyValThrArgHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgVal 332
   |||
Db 1582 TATAGAGAGTGACACAGCACCATCAGCATGGAAGTGGCAGGCAAGATTTGGAAGAGTT 1641
   |||
Qy 333 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaGluAla 352
   |||
Db 1642 GCAGGGAACAAAGATCTCTATCTTGGGACTTTCAGCACACAGAAAGAGCTGCGAAGCA 1701
   |||
Qy 353 TyrAspIleAlaAlaLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsn 372
   |||
Db 1702 TATGATGTGTGCAATCAAAATTCAGAGGTGAATGTGTCAACACTTTGACATATCG 1761
   |||
Qy 373 ArgTyrAsnValLysAlaIleLeuGluSerProSerLeuProIleGlySerAlaLys 392
   |||
Db 1762 CGATACGAGCTGGGAAATCATGGCTAGTAATACACTTCTCTGCTGGAGAAATAGTAGG 1821
   |||
Qy 393 ArgLeuLysGluAlaAsnArgProValProSerMetMetMetIleSerAsnAsnValSer 412
   |||
Db 1822 AGAAACAAAGAAAGAGAG-----CCAAATGAATACAACTCTCAGT 1860
   |||
Qy 413 GluSerGluAsnSerAlaSerGlyTyrGlnAsnAlaAlaValGlnHisGlnGlyVal 432
   |||
Db 1861 GATCACAGAATGAAGAGCATGTTTCAAAATGACACAC---AACACACAAATGGGAAT 1917
   |||
Qy 433 AspLeuSer-----LeuLeuHisGlnHisGlnGluArgTyrAsnGlyTyrTyr 448
   |||
Db 1918 AATGTCACAGATTGGAAAATGTTGTCATATCAATCTCCAAATCTCTCTAGGAATAT 1977
   |||
Qy 449 TyrAsn-----GlyGlyAsnLeuSerSerGlu 457
   |||
Db 1978 AGAALCCCAACTTTTCCATGTCATTACAAGATCTAATCGGTATTGATTGATGATCA 2037
   |||
Qy 458 SerAlaArgAlaCysPheLysGlnGluAspGln-----HisHisPhe 472
   |||
```


135	LeuSerMetIleIysThrTrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGlu	154
306	-----CCAGACGACGGCCGACGCGCGCA	332
155	Asn-----	155
333	GCTGTAGAGCTCGAGGCTCAAGTTTCGCGCGCGGCTTCCAGCTGAGCGCGCGCTGG	392
156	GlyAsnAlaAlaIysGlyLeuSerLeuSerMetAsnSerThrSerCysAspAsnAsn	175
393	GGCGGGCGCCACGAGTCGCGCGTGTTCGCGCGGCTGCTTAGAGCA-----GACCGA	446
176	AsnAspSerAsnAsnValAlaGlnGlyLysThrIleAspAspSerValGluAla	195
447	CGAGACGACGAGCTGCGCTTGGCGCGCAGGCGCG-----GTTGGCGCGCGCGC	497
196	ThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValTh	215
498	GGAGCAGAGAAGCGGTGGACTCGTTCGGGCAGAGAACGTCTATCTACCGCGCGCTCAC	557
215	ArgHisArgTrpThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgG	235
558	ACGGCATCGGTGACCGGAAGGTACAGGCGCATCTGTGGGATTAACAGCTGCCCGCGTGA	617
235	uGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGluLysAl	255
618	GGGCCAGAGCGCGCAAGGACCCAAAGTGTTCTGGTGGCTATGATAAGAGAGAGAAAGC	677
255	AlaArgAlaTyrAspLeuAlaIleuLysTyrTrpGlyThrThrThrThrThrAsnPh	275
678	GGCGAGGCGGTATGATCTTCGCGGTTGAAGTACTGGGGTCTTAGACCAACACCAACT	737
275	eProMetSerGluTyrGluLysGluValGluMetLysHisMetThrArgGlnGluTyr	295
738	TCGGTTCGCGAGTACGAAAGGAGCTTGAGGAGATCAAGCATATGACAGCCAGAGTT	797
295	rValAlaSerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgG	315
798	TGTAGCTTCTTAGGAGGAGAGACAGTGGATTCTCTCGTGGTCTTCTATATACAGGG	857
315	yValThrArgHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAs	335
858	TGTTTCAAGACATCACCAGCATGGCGCATGGCAAGCAAGATCGAAGATGGCTGGCAA	917
335	nLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaGluAlaTyrAspI	355
918	CAAGGATCTCTACCTCGAACATTTCGCCACTGAGGAGGAAGCTGCAGAAGCTTATGACAT	977
355	eAlaAlaIleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAs	375
978	CGCGCGCATCAAGTTCGTCGCTGGAATGTGTGCCAACTTTGAGATCGCGCGGTATAA	1037
375	nValLysAlaIleLeuGluSerProSerLeuProIleGlySer---AlaAlaLysArgLe	394
1038	CGTGGAGAGCATATCAGTAGC---AACCTTCCAATCGGAGCATGCCCGCAACCGAG	1094
394	uLysGluAlaAsn-----ArgProValProSerMetMetIle-----	407
1095	CACCAAGGCGCGCTAGAGCTCGCTCCGTCCAGCTCGCGCGACCCATCGCGCACCGA	1154
408	SerAsnAsnValSerGluSerGluAsnSerAlaSerGlyTyrGlnAsnAlaValG	427
1155	AGCGAACACACCGCGGTGCGGCCACCGTCGACCGCTTCGCTTCACAGCTCTTCCAATGA	1214
427	nHisGlnGlnGlyValAspLeuSerLeuLeuHis---GlnHisGlnGluArgTyrAsn--	445
1215	GTATGATCAGGCAGATTACCTGTCTATACCTCGCATTCGAGCATCACCAGCAAGAAACCT	1274
446	-----GlyTyrTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaAr	460
1275	TCAGGGGTAGGCTTCGCTCTCTACAGCTCAGGTGAACCTGGATTTCGCAATGCGAA	1334
460	qAla-----CysPhe-----LysGlnGluAs	467

Db	1335	TGCAATGAGGCATCTCAAAATTCCTACCAATGCTCTATTACATGACGAGCAGCA	1394
Qy	467	pAspGlnHisHisPheLeuSerAsnThrGlnSerLeuMetThrAsnIleAspHisGlnSe	487
Db	1395	ACATCAGCACCATCAGCAGCAGCAGCAGCAA	1439
Qy	487	rSerValSerAspAspSerValThrValCysGlyAsnValValGlyTyTGlyTyG	507
Db	1440	TGATCAGTCAAGAAGCTCCAAACACAGCTGCGGCTCCAT	1479
Qy	507	nGlyPheAlaAlaProVal	513
Db	1480	CGTTTGCAACACCCATCAGTCTTTCAGTGGGTCTATATGAAGCTCCATGACGACGAGCAGG	1538
Qy	514	-----AsnCysAspAlaTyAlaAlaSerGluPhe-AspTyAsnA	527
Db	1539	CACCTTTGGGATACTACCAAAATGTTGCAGCCTTTCAGACACCAATCTTTGGGATGGATG	1598
Qy	527	laArgAsnHisTyTyPheAlaGlnGlnGlnThrGlnGlnSerProGlyGlyAspp	547
Db	1599	ATTGGGGAGGTGGCACTTTGCAGGCAAAAGGATGAAGAGCTGATCCTCGATCAGAGG	1658
Qy	547	heProAlaAlaMetThrAsnAsnValGlySerAsn	558
Db	1659	TGATGAGGATCAAGACTGACATAGTTCATCTTCAC	1693
RESULT 12			
AX555220		1926 bp	DNA linear PAT 27-NOV-2002
LOCUS	AX555220	Sequence 5 from Patent WO02059332.	
DEFINITION	AX555220		
ACCESSION	AX555220		
VERSION	AX555220.1	GI:25898748	
KEYWORDS			
SOURCE	Oryza sativa		
ORGANISM	Oryza sativa		
REFERENCE			
AUTHORS	He, S. and Dotson, S.B.		
TITLE	Nucleic acid molecules associated with plant cell proliferation and growth and uses thereof		
JOURNAL	Patent: WO 02059332-A 5 01-AUG-2002;		
FEATURES	Monsanto Technology LLC (US)		
source	Location/Qualifiers		
	1..1926		
	/organism="Oryza sativa"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:4530"		
CDS	1..>1926		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD58227.1"		
	/db_xref="GI:25898748"		
	/translation="MASGGGGSNWLGSLSFHPAMEVPSRSTAAHHHHHPAPAAH AAAAGASPPPSATTCNLFSPFAAQVAPSGYVVGAGDGDSTAGVYSHLE VMPKSDGSLIMEGMPSSPKLEDFLGGNGSHDPATYYISQQAEDASRAYOH HQLVPYNYFLTAEMLEQAAAPMEDAMAAKNFLVTSYGACNGEMFQPLSLMS PGSSSCVAAAPQHQHQAFAAAAGAGCGGNSNDGGEQVKGKRGTKGQKQFDP VHRKSIDTFGQYVSRVTHRWGTYEARHWDNSCKDQGTAKGRVYLGVDYDE KAARYDALAKYGLSTHIFPLENVRDEIEEMERTROYVAHLRRRSRSGFSGAS LYRGVTHHGHGWAQISGVAGNKDLYLGTFTQEEAAEAYDIAATKFGNLAVTNF DITRYVDKTMESSLLPGAARVKAIEADPHVPIGRELGAESAATVTCDDMR MWLHGSQQQAACCTEATADLQKGFMDGDAHSLHGVGFVDESAAAEIDVPGGKISG INTNSSSLVTSLSNREGSPERLGLANLYAKKHPTAVSLAAMFMFMFAPAAAAH RPPSAIAHLFVFAWTDTA"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1,28e-60	Length:	1926
Score:	973.150	Matches:	254

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE AUTHORS

The Rice Full-Length cDNA Consortium, National Institute of
Agricultural Sciences Rice Full-Length cDNA Project Team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yabuchi, J., Ishikawa, M., Yamada, H., Coka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yabuchi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group, Ohtsuki, K., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Mura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Ohtsuki, K., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Ohtsuki, T.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
Japanese rice
Science 301 (5631), 376-379 (2003)
12869764

2 (bases 1 to 2510)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koyama, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Otsu, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagami, A., Takahashi, F., Takaku-Akashira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yabuchi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yabuchi, J., Yokomizo, S. and
Yoshimura, A.

Direct Submission

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agricultural Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp)
Tel: 81-29-838-7007, Fax: 81-29-838-7007
This clone is one of the 28K full-length cDNA clones from japonica
rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team, Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yabuchi, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yabuchi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayatsu, N., Imotani, K., Hayatsu, N.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J.,

FEATURES source

1..2510
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/db_xref="taxon:39947"
/clone="002-101-C10"

ORIGIN

Alignment Scores:
Pred. No.: 1.77e-60 Length: 2510
Score: 973.50 Matches: 254
Percent Similarity: 50.48% Conservative: 63
Best Local Similarity: 40.45% Mismatches: 155
Query Match: 31.21% Indels: 158
DB: 8 Gaps: 19

US-09-980-364-2 (1-579) x AK106306 (1-2510)

QY 2 AsnAsnAsnTrpLeuGlyPheSerLeuSerPro----- 12
DB 200 AGCAGCAACTGGTAGGCTTCTCGCTCTCCCGCACATGCGCGCATGGAGTGGCGCTCC 259
QY 13 -----TyGluGlnAsnHisArgLysAspValTyrSer 24
DB 260 TCCTCTGAGCATGACGTGCTGCTCATCATCATCATCATCATCATCATCATCATCATCAT 319
QY 25 SerThrThrThrValValAspValAlaGlyGluTyrCysTyrAspProThrAlaAla 44
DB 320 GCT-----GCTGCGGAGCCATGCTGCTCTCCCGACAGCGCC 358
QY 45 Serasp-----GluSerSerAlaIleGlnThrSerPheProSerProPhe 59
DB 359 AGCAGCTCAACT 418
QY 60 GlyValValValAlaPhePheThrArgAspAsnAsnSerHisSerArgAspTrp----- 77
DB 419 TACTACTACGTGCGGCG 478
QY 78 -----AspIleAsnGlyCysAlaCysAsnAsnIleHisAsnAsp 90
DB 479 CACCT 538
QY 91 GluGlnAspGlyProLysLeuGluAsnPheLeuGly-----Arg 103
DB 539 CGTCTCATCGCCCAAGCTCGAGGACTTCTTGGGTGTGGCAATGGCAGTGGCGCATGAC 598
QY 104 Thr 119
DB 599 CGGCGCCACTACTATAGCCAGCGGCCAAGAGCAGCAGGATGCAAGCAG-GGGCGGCTACCA 657
QY 119 ----- 119
DB 658 GCACCCAGCAGTAGTCTCCCTCAACTACCAGCCCATTCACGGAAGCAGAGATGCTGCAAGA 717
QY 120 -----SerGlyCysTyrGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 137
DB 718 GCGCGCAGCGCGCGCAATGGAGGACGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 777
QY 138 IleIleYThrThrLeuLeuArgGlnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsn 157
DB 778 CTACGCGCGCTGCTACGG-----CAACCA-----GGAGAT 807
QY 158 AlaAlaLysGlyLeuSerLeuSerMet-----AsnSerSerThrSerCys----- 172

Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itch, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurotsaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

TITLE JOURNAL

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, I., Arakawa, I., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itch, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1. 1745
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clones="J033075007"

ORIGIN

Alignment Scores:

Pred. No.:	6,42e-60	Length:	1745
Score:	963.00	Matches:	233
Percent Similarity:	54.66%	Conservative:	66
Best Local Similarity:	42.60%	Mismatches:	122
Query Match:	30.88%	Indels:	126
DB:	8	Gaps:	19

US-09-980-364-2 (1-579) x AK101959 (1-1745)

Qy	25	SerThrThrThrThrValValValValAlaGlyGluTyrCysTyrAspProThrAlaAla	44
Db	2	AGCACAACTGCTTACATCGAT---CAATAGCTTTTAGTATAT---CTTCCCTCTCTCA	55
Qy	45	SerAspGluSerSerAlaGlnThrSerPheProSerPheGlyValValValAsp	64
Db	56	TAAGCTCAAAACAAACGCTGGTCAAGTCTCTTCAAGATCGGCGCATCATATCTATC	115
Qy	65	AlaPheThrArgAsp---AsnAsnSerHisSerArgAspTyrAspIleAsnGlyCysAla	83
Db	116	GCCATGCATGCATGCATGCATGCATGCATGCATGCATGCATGCATGCATGCATGC	172
Qy	84	CysAsnAsnIleHisAsn-----	89
Db	173	TCTTAACAACATACCAACATACCAACATACCAACATACCAACATACCAACATACCA	232
Qy	90	-----AspGluGlnAspGlyProLysLeuGluAsnPheLeuGlyArgThrThr	105
Db	233	GGAGAGAGGGGGCCAGCGGGCGCTCCGAAGATGGAGGATTTCTC-----	280
Qy	106	ThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGlySerGlyCysTyrGlyGly	125
Db	281	-----GGCGGC-----	286
Qy	126	GlyAspGlyGlyGlyGlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGln	145
Db	287	CTAGCGGAGGGGGCGGGCGCTCCGCGCGCTCCGCGAGCTGCC-----	331
Qy	146	ProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaAlaLysGlyLeuSerLeuSer	165
Db	332	CCGAGAGATCAGCTCAGCTGCGCGGAGCTGTGTAGCATCGCGCGGGTTC-----	382
Qy	166	MetAsnSerSerThrSerCysAspAsnAsnAspSerAsnAsnValAlaGln	185
Db	383	TTGCGCGGTATCCCGCGCTCGAAGC-----GGCGGC-----	415
Qy	186	GlyLysThrIleAspAspSerValGluAlaThr-----ProLysLys	199
Db	416	GGGTGACCATCGCATCGGACCGACCGCGGGGGAGCTGGCGATCCCGCGGAGG	475
Qy	200	ThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisArgTrp	219
Db	476	ACCGCGGAGACGCTTCGGCACCGACGCTCATCTACCGTGTGTGGCTAGGACCGGTGG	535
Qy	220	ThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGlyGlnThrArg	239
Db	536	ACGGGGAGGTAGAGCGGCACCTGTGGACCAATAGCTCCCGCGGAGGCGAAGCGCGC	595
Qy	240	LysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyr	259
Db	596	AAAGCCCGCCAGTCTACTTAGGAGGTATGATAGGAGGAGAGGCGCGCAAGAGCTTAC	655
Qy	260	AspLeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrThrThrThrThrThrThr	279
Db	656	GGCTTCGCGCGCTTAAAGTCTGGGGTCCACCCACGACGAACTTCCCTGTGTCCCAAC	715
Qy	280	TyrGluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeu	299
Db	716	TACGAGACGGAGCTGGAGAGATGAATGCCATGACCGCGGAGGAGTTCATCGCTCGCTG	775
Qy	300	ArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHis	319
Db	776	CGCAGGAAGACGCGGCTTCTCAAGAGGGGCTTCCATCTACAGAGGAGTAACAGACAT	835
Qy	320	HisGlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyr	339
Db	836	CATCAGCAGCGCGGTGGCAAGCGGAGTTCGCGAGGTGCGCGAACAAGAGCTCTGAC	895
Qy	340	LeuGlyThrPheGlyThrGlnGluGluAlaGluAlaTyrAspIleAlaAlaIleLys	359
Db	896	TTGGSCATTTTCAGCAGCAGGAGGCTGCCAGGCGTACGACATTCGCGCTATCAAG	955
Qy	360	PheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIle	379

GenCore version 5.1.6
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 19, 2004, 10:20:12 ; Search time 653 Seconds
(without alignments)

3766.779 Million cell updates/sec

US-09-980-364-2

Title:

Perfect score: 3119

Sequence: 1 MNNNLGFSLSPEQNHRK.....YYHGGGGEVAPFTFWNDN 579

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US0980364/runat_15072004_095235_25858/app_query.fasta_1.775
-DB=N_Geneseq_29Jan04 -OFM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOCPCL=0
-LOPEXT=0 -UNITS=bits START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=500 -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-US=US0980364@cgn_1.1_470 @runat_15072004_095235_25858 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3119	100.0	2014	4	AAc91398 Brassica
2	3031	97.2	2011	4	AAc91399 Brassica
3	2690.5	86.3	4873	4	AAc91400 Brassica
4	1310	42.0	2588	7	AAc00801 Glycine m
5	1296	41.6	2463	7	AAc00803 Glycine m
6	1119.5	35.9	1990	7	AAc00806 Glycine m
7	1075.5	34.5	1824	7	AAc00802 Glycine m
8	1047.5	33.6	1818	9	ADD30878 Plant yie

9	1047.5	33.6	1818	9	ADD30822	ADD30822	Plant yie
10	1047.5	33.6	1818	9	ADD30333	ADD30333	Plant yie
11	1017	32.6	1935	7	ACC00848	ACC00848	Gossypium
12	1003	32.2	1959	7	ACC00847	ACC00847	Catalpa s
13	1000	32.1	1878	7	ACC00846	ACC00846	Glycine m
14	996.5	31.9	1954	7	ACC00850	ACC00850	Glycine m
15	987.5	31.7	1874	7	ACC00799	ACC00799	Glycine m
16	979	31.4	2374	7	ACC00849	ACC00849	Zea mays
17	977	31.3	2052	9	ADD30380	ADD30380	Plant yie
18	974.5	31.2	1941	9	ADD30349	ADD30349	Plant yie
19	973.5	31.2	1926	6	ABN84482	ABN84482	Rice AINT
20	971	31.1	2168	7	ACC00804	ACC00804	Glycine m
21	964.5	30.9	1758	7	ACC00794	ACC00794	Oryza sat
22	957.5	30.7	1290	7	ACC00805	ACC00805	Glycine m
23	948.5	30.4	1753	7	ACC00792	ACC00792	Zea mays
24	948	30.4	1864	7	ACC00791	ACC00791	Zea mays
25	942	30.2	2010	6	ABN84484	ABN84484	Rice AINT
26	942	30.2	2344	6	ABN84480	ABN84480	Soybean A
27	941	30.2	1500	9	ADD30559	ADD30559	Plant yie
28	924	29.6	2323	6	ABN84481	ABN84481	Soybean A
29	913.5	29.3	1738	3	AAAS9222	AAAS9222	CDNA enco
30	891	28.6	1668	3	AAAS9222	AAAS9222	Arabidops
31	891	28.6	2148	3	AAAS9220	AAAS9220	CDNA enco
32	891	28.6	2148	7	ABX13437	ABX13437	A. thalia
33	771	24.7	1749	4	AAc83248	AAc83248	Cellulose
34	714	22.9	1271	7	ACC00797	ACC00797	Glycine m
35	711.5	22.8	1619	7	ACC00843	ACC00843	Momordica
36	673	21.6	1726	7	ACC00845	ACC00845	Glycine m
37	672	21.5	1710	7	ACC00785	ACC00785	Zea mays
38	671	21.5	551	7	ACC00796	ACC00796	Oryza sat
39	670	21.5	1065	9	ADD31036	ADD31036	Plant yie
40	670	21.5	1065	9	ABE31572	ABE31572	Plant yie
41	667	21.4	1753	7	ACC00844	ACC00844	Ricinus c
42	657	21.1	999	3	AGC51369	AGC51369	Arabidops
43	652.5	20.9	1879	7	ACC00789	ACC00789	Zea mays
44	647.5	20.8	1465	6	ABQ81396	ABQ81396	Arabidops
45	647.5	20.8	1530	6	ABK65225	ABK65225	Arabidops

ALIGNMENTS

RESULT 1

AAc91398

ID AAC91398 standard; cdna; 2014 BP.

XX AC AAC91398;

XX AC AAC91398;

DT 19-MAR-2001 (first entry)

XX DE Brassica napus BNM3A cdna.

XX DE Brassica napus BNM3A cdna.

XX KW Brassica napus microspore embryo; BNM; microspore embryogenesis; BNM3A;

XX KW asexually derived embryo production; increased regenerative capacity; ss.

XX OS Brassica napus.

XX OS Brassica napus.

XX PN EP1057891-A1.

XX PD 06-DEC-2000.

XX PF 02-JUN-1999; 99EP-00201745.

XX PF 02-JUN-1999; 99EP-00201745.

XX PR 02-JUN-1999; 99EP-00201745.

XX PR 02-JUN-1999; 99EP-00201745.

XX PA (CPRO-) CPRO-DLO CENT PLANTENVEREDELINGS REPROD.

XX PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

XX PI Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;

XX PI Boutillier K;

XX XX WPI; 2001-026720/04.

XX DR P-PSDB; AAB50872.

XX DR P-PSDB; AAB50872.

PT Brassica napus BNM3 DNA sequences, useful for transforming host cells to
 PT produce asexually derived embryos.

XX Claim 4; Fig 2; Sipp; English.

XX The present sequence is given in a specification relating to a Brassica
 CC napus microspore embryo (BME) gene obtained during the induction of
 CC microspore embryogenesis. The isolated DNA molecule can be used in a
 CC vector that is used to transform plant cells. The vector can be used in
 CC methods to produce asexually derived embryos, for modifying the
 CC regenerative capacity of a plant, and for producing an apomictic plant

XX Sequence 2014 BP; 596 A; 411 C; 462 G; 545 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,28e-268 Length: 2014
 Score: 3119.00 Matches: 579
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-980-364-2 (1-579) x AAC91399 (1-2014)

QY 1 MetAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLys 20
 DB 120 ATGAATAATACTGGTTAGGCTTTTCTCTCTCTTATGACAAATACCATCGTAAG 179
 QY 21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
 DB 180 GACGCTCTACTCTTCACCAACACACACACCTGCTAGATGCGCGGAGAGTACTGTTCGAT 239
 QY 41 ProThrAlaLeuSerAspGluSerSerAlaLeuGlnThrSerPheProSerProPheGly 60
 DB 240 CCGACCGCTGCTCCGATGAGTCTTCAGCCATCCAAACATCGTTTCTCTCCCTTTGGT 299
 QY 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAspIleAsn 80
 DB 300 GTCGTCGTCGATGCTTTTACCAGAGACAAACATAGTCACTCCCGAGATGGGACATCAAT 359
 QY 81 GlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsnPhe 100
 DB 360 GGTTCGTGCATGCAATACATCCACACGATGAGCAAGATGACCAAGCTTGAGAAATTC 419
 QY 101 LeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGlySer 120
 DB 420 CTGCGCCGACACACACGATTTACACACCAACGAAACGTTGGAGATGGAAATGGAAGT 479
 QY 121 GlyCysTyrGlyGlyGlyGlyGlyGlyGlySerLeuGlyLeuSerMetIleLysThr 140
 DB 480 GCGTGTATGAGGAGGAGACGGTGTGTGTGTCTACTAGGACTTTCGATGATAAGACA 539
 QY 141 TrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaLys 160
 DB 540 TGGCTGAGAAATCAACCGCTGGAATGTTGATATCAAGAAATGCGCAATGCTGCAAAA 599
 QY 161 GlyLeuSerLeuSerMetAsnSerThrSerCysAspAsnAsnAsnAspSerAsnAsn 180
 DB 600 GGCCTGTCCTCTCAATGAATCACTACTCTTCTGTGATACACACACGACGACGATAC 659
 QY 181 AsnValValAlaGlnGlyLysThrIleAspAspSerValGluAlaThrProLysLysThr 200
 DB 660 AACGTTGTTGCCCCAAGGAGAGACTATTGATGATAGCGTTGAAAGCTACACCGAAGAAACT 719
 QY 201 IleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisArgTyrThr 220
 DB 720 ATTGAGAGTTTGGACACAGAGAGCTCTATATACCGCGGTGTTACAGGATCCGCTGACA 779
 QY 221 GlyArgTyrGluAlaHisLeuTyrAspAsnSerCysLysArgGluGlyGlnThrArgLys 240
 DB 780 GGAAGATATGAGGCACATTTATGGATAATAGTTGTAAGAGAGAGGCCCAACCGCCAAA 839
 QY 241 GlyArgGlnValTyrLeuGlyGlyTyrAspLysGluLysAlaAlaArgAlaTyrAsp 260

DB 840 GGAAGCAAGTTTATTGGAGGTTTATGACAAAGAAAAAGACAGCTAGGCGCTTATGAT 899
 QY 261 LeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrAsnPheProMetSerGluTyr 280
 DB 900 TTAGCCGCACTCAAGTATTGGGAACCACTACTACTACTTCCCATGAGCGAATAT 959
 QY 281 GluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArg 300
 DB 960 GAAAAAGAGGTGAAGAGATGAAGCAGCATGACAAAGCAAGAGTATGTTGCCCTCACTGCC 1019
 QY 301 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHis 320
 DB 1020 AGGAAAGTAGTGGTTCTCTCGTGTGCTGATGATTTATCTGGAGTAAACAGACATCAC 1079
 QY 321 GlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeu 340
 DB 1080 CAACATGGAAGATGGCAAGCTAGGATAGGAGAGTCCCGGTAAACAAAGACCTCTACTTG 1139
 QY 341 GlyThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAspIleAlaAlaLysPhe 360
 DB 1140 GGAACCTTTGGCACACAAAGAGAGCTGCGAGGAGATACGACATTCGGGCCATCAATTC 1199
 QY 361 ArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeu 380
 DB 1200 AGAGGATTAACCGCAGTGAATTCGACATGAACAGACATACAGTTAAAGCAATCCTC 1259
 QY 381 GluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgPro 400
 DB 1260 GAAAGCCCTAGTCTTCTTATTGTTAGCGCGCAAAACGTTCTCAAGAGAGCTAAACGTCG 1319
 QY 401 ValProSerMetMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerGly 420
 DB 1320 GTTCCAAGTATGATGATGATCACTAGTAATTAACGTTTCAGAGAGTGAAGTAGTCTAGCG 1379
 QY 421 TrpGlnAsnAlaAlaValGlnHisHisGlnGlyValAspLeuSerLeuHisGlnHis 440
 DB 1380 TGGCAAAACGCTCGCGGTTTCAGCATCATCAGGAGTAGATTTTGGAGTTATTGCCAACAT 1439
 QY 441 GlnGluArgTyrAsnGlyTyrTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaArg 460
 DB 1440 CAAGAGAGGTACAATGGTTTATTATTAATGAGAGAAACTTGTCTTCGAGAGTGTAGG 1499
 QY 461 AlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeuMet 480
 DB 1500 GCTTGTGTTCAAAACAGAGAGATGATCAACACCATTTCTTTGAGCAACACGACGAGCTCATG 1559
 QY 481 ThrAsnIleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsnVal 500
 DB 1560 ACTAATATCGATCATCAAGTCTGTTTCGGATGATTCGGTTACTGTTTGGGAAATGTT 1619
 QY 501 ValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAspAlaTyrAlaAla 520
 DB 1620 GTTGGTTATGGTGGTTATCAAGAGATTTTCAGCCCGGTTAACTGCGATCGCTCGCTGCT 1679
 QY 521 SerGluPheAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGln 540
 DB 1680 AGTGAGTTTGATTAACGCAAGAAACCATTTACTTTTGTCTCAGCAGCAGCAGCCGAG 1739
 QY 541 GlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyr 560
 DB 1740 CAGTCGCCAGGTGGAGATTTTCCCGCGCAATGACGAATAATGTGGCTCTTAATATGAT 1799
 QY 561 TyrHisGlyGluGlyGlyGlyGluValAlaProThrPheThrValTyrAsnAsnAsn 579
 DB 1800 TACCATGGGAAGGTGGTGGAGAAAGTTGCTCCAACATTTACAGTTTGGACGACAT 1856

RESULT 2

AAC91399

ID AAC91399 standard; cDNA; 2011 BP.

XX AAC91399;

AC AAC91399;

XX

DT 19-MAR-2001 (first entry)
 XX Brassica napus BNM3B cDNA.
 DE
 XX
 KW Brassica napus microspore embryo; BNM; microspore embryogenesis; BNM3B;
 KW asexually derived embryo production; increased regenerative capacity; ss.
 XX
 OS Brassica napus.
 XX
 EP1057891-A1.
 PN
 XX
 XX 06-DEC-2000.
 PD
 XX
 XX 02-JUN-1999; 99EP-00201745.
 PF
 XX
 XX 02-JUN-1999; 99EP-00201745.
 PR
 XX
 XX (CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
 PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
 PA
 XX Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;
 PI Boutillier K;
 PI
 XX
 DR WPI; 2001-026720/04.
 DR P-PSDB; AAB50873.
 XX
 PT Brassica napus BNM3 DNA sequences, useful for transforming host cells to
 PT produce asexually derived embryos.
 XX
 XX Claim 4; Fig 2; 51pp; English.
 PS
 XX
 CC The present sequence is given in a specification relating to a Brassica
 CC napus microspore embryo (BNM) gene obtained during the induction of
 CC microspore embryogenesis. The isolated DNA molecule can be used in a
 CC vector that is used to transform plant cells. The vector can be used in
 CC methods to produce asexually derived embryos, for modifying the
 CC regenerative capacity of a plant, and for producing an apomictic plant
 CC
 XX SQ Sequence 2011 BP; 610 A; 395 C; 462 G; 544 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3,67e-260 Length: 2011
 Score: 3031.00 Matches: 564
 Percent Similarity: 98.27% Conservative: 5
 Best Local Similarity: 97.41% Mismatches: 10
 Query Match: 97.18% Indels: 0
 DB: 4 Gaps: 0
 US-09-980-364-2 (1-579) x AAC91399 (1-2011)
 Qy 1 MetAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisHisArgLys 20
 Db 51 ATGAATAAATACCTGGTTAGGCTTTTCTCTCTCTATGAACAAATCACCATCGTAAG 110
 Qy 21 AspValTyrSerSerThrThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
 Db 111 GACCTCTGCTCTCCACACACACACCGCGGTAGATGTCGCGGAGAGTACTCTTAGAT 170
 Qy 41 ProThrAlaAlaSerAspGluSerSerAlaAlaGlnThrSerPheProSerProPheGly 60
 Db 171 CCGACCGCTGCTCCGATGAGTCTTCAGCATCCAAACATCGTTTCTCTCTCTTGGT 230
 Qy 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAspIleAsn 80
 Db 231 GTCGCTCTCGATGCTTTCCACAGACACAAATAGTCACTCCCGAGATTGGGACATCAAT 290
 Qy 81 GlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsnPhe 100
 Db 291 GGTAGTGCATGTATATACATCCCAATGATGAGCAAGATGAGCCAAACTTGAGAAATTC 350
 Qy 101 LeuGlyArgThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 120
 Db 351 CTTGCCCGCACCCACCATGATTACAAACACCAACCAACCAACCAACCAACCAACCAACCAAC 410

121 GlyCysTyrGlyGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMetIleLysThr 140
 411 GGGTGTATGGAGGAGGAGCGGTGGTGGTCTACTAGGACTTTCGATGATAAGACA 470
 141 TrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaLys 160
 471 TGGCTGAGAANTCAACCCGCGGTAATGTTGATAATCAAGAAATGCAATGGTGCAAA 530
 161 GlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAsnAspSerAsn 180
 531 GGCCTGTCCCTCTCAATGAACCTACTTCTTGTGATAACAACAACCTACAGAGTAAC 590
 181 AsnValValAlaGlnGlyLysThrIleAspAspSerValGluAlaThrProLysLysThr 200
 591 AACCTTGTGCCCAAGGAGAGCTATTGATGATAGCGTGGAGGTACACCGAAGAACT 650
 201 IleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisArgTrpThr 220
 651 ATTGAGAGTTTTGGACAGAGGACGCTATATACCGCGGTGTTCAAGGCGATCGGTGACA 710
 221 GlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArgGluGlyGlnThrArgLys 240
 711 GGAGATATAGGCGACATTTATGGGATATAGTTGTAACGAGAGGCGCAACCGCGCAA 770
 241 GlyArgGlnValTyrLeuGlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyrAsp 260
 771 GGAACACAAGTTTATTTGGGAGGTTATGACAAAGAGAAAAGCAGCTAGGCGTTATGAT 830
 261 LeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrThrThrThrThrThrThrThr 280
 831 TTAGCGGCACTCAAGTATTGGGAAACCCACCTACTACTAACTTCCCATGAGCGAATAT 890
 281 GluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArg 300
 891 GAGAAAGAGATAGAGAGATGAGCAGCATGACAGGCAAGAGTAGTTGCTCCTCACTCGC 950
 301 ArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisHis 320
 951 AGGAAAAGTAGTGGTTTCTCTCGTGCATCGATTTATCGTGGAGTAAACAGACATCAC 1010
 321 GlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeu 340
 1011 CAACATGGAGATGGCAAGCTAGGATAGAGAGTCCGCGGTAAACAAGAGCTCTACTTGG 1070
 341 GlyThrPheGlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaLysPhe 360
 1071 GGAACCTTTGGCACACAAGAGAGCTGACAGGCGCATACGACATTCGCGCCATCAAAATTC 1130
 361 ArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValIleAlaLysLeu 380
 1131 AGAGGATTAAACCGCAGTGACTTAACCTGACATGAACAGATACACGTTAAAGCAATCCCTC 1190
 381 GluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgPro 400
 1191 GAAAGCCCTAGTCTTCTCTTCTTCTGTCGCGCGCAAAACGCTCTCAAGGAGGCTAACCGTCCG 1250
 401 ValProSerMetMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerGly 420
 1251 GTTCCAAGTATCATGATCATGATCAATACGTTTCAGAGAGTGAGAAATATCTAGCGGT 1310
 421 TrpGlnAsnAlaAlaValGlnHisHisGlnGlyValAspLeuSerLeuHisGlnHis 440
 1311 TGGCAAAACCTCGCGTTCAGCATCATCAGGAGTAGATTGAGCTTATTGCGAGCAACAT 1370
 441 GlnGluArgTyrAsnGlyTyrTyrTrpAsnGlyGlyAsnLeuSerSerGluSerAlaArg 460
 1371 CAAGAGAGTACATGTTTATTATTAACAATGAGGAAATCTGCTCTCGAGAGTCTAGG 1430
 461 AlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeuMet 480
 1431 GCTTGTTCACAAACAGAGGATGATCAACCCATTTCTTGAGCAACACCGCAGGCTCATG 1490

QY 481 ThrAsnIleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsnVal 500
 Db 1491 ACTAATATCGATCATCAAGATTCTGTTTCAGATGATTCGGTTACTGTTTGTGGAATGTT 1550
 QY 501 ValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAspAlaTyrAlaAla 520
 Db 1551 GTTGTTTATGTTGTTATCAGGATTTCAGCCCGGTTAACTGCGATGCGCTACGCTGCT 1610
 QY 521 SerGluPheAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGln 540
 Db 1611 AGTGAGTTTGACTATAACGCAAGAAACCAATTACTTGTCTCAGCAGCAGCAGCCAG 1670
 QY 541 GlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyr 560
 Db 1671 CATTCCGAGGAGGAGATTTCCTCCCGCAATGACGAATAATGTTGGCTCTAATATGAT 1730
 QY 561 TyrHisGlyGlyGlyGlyGlyGluValAlaProThrPheThrValTyrAsnAspAsn 579
 Db 1731 TACCATGGGAAGTGGTGGAGAGTTGCTCCAACTTACAGTTTGGAAACGCAAT 1787

RESULT 3

AAC91400
 ID AAC91400 standard; DNA; 4873 BP.
 AC AAC91400;
 DT 19-MAR-2001 (first entry)
 XX
 DE Brassica napus BMN3A DNA sequence.
 XX
 KW Brassica napus microspore embryo; BMN; microspore embryogenesis; BMN3A;
 KW asexually derived embryo production; increased regenerative capacity; ds.
 OS Brassica napus.
 XX
 PN EP1057891-Al.
 XX
 PD 06-DEC-2000.
 XX
 PF 02-JUN-1999; 99EP-00201745.
 XX
 PR 02-JUN-1999; 99EP-00201745.
 XX
 PA (CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
 PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
 PI Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;
 PI Boutilier K;
 XX
 DR WPI; 2001-026720/04.
 DR P-PSDB; AAB50872.
 XX
 PT Brassica napus BMN3 DNA sequences, useful for transforming host cells to
 PT produce asexually derived embryos.
 XX
 PS Claim 1; Page 32-33; 51pp; English.

CC The present sequence is given in a specification relating to a Brassica
 CC napus microspore embryo (BMN) gene obtained during the induction of
 CC microspore embryogenesis. The isolated DNA molecule can be used in a
 CC vector that is used to transform plant cells. The vector can be used in
 CC methods to produce asexually derived embryos, for modifying the
 CC regenerative capacity of a plant, and for producing an apomictic plant
 XX
 SQ Sequence 4873 BP; 1596 A; 856 C; 907 G; 1514 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,6e-229 Length: 4873
 Score: 2690,50 Matches: 574
 Percent Similarity: 53.24% Conservative: 1
 Best Local Similarity: 53.15% Mismatches: 4
 Query Match: 86.26% Indels: 505
 Gaps: 8

US-09-980-364-2 (1-579) x AAC91400 (1-4873)
 QY 1 MetAsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisHisArgLys 20
 Db 1620 ATGAATATATACCTGGTTAGGCTTTCTCTCTCTCTATGAACAAATACCATCGTAAAG 1679
 QY 21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGlyTyrCystTyrAsp 40
 Db 1680 GAGCTCTACTCTCCACCACCAACCGTCGTAGATGTCGCGGAGAGTACTGTATCCGAT 1739
 QY 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
 Db 1740 CCGACCGCTGCTCCGATGAGTCCTTCAGCCATCCAAACATCGTTTCTCTCCCTTTGGT 1799
 QY 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArg----- 75
 Db 1800 GTCGTCGTGATGCTTTCACCAGACACAACTAGTCACTCCCGAGGTTATTGTTTGA 1859
 QY 75 ----- 75
 Db 1860 ACTACTGTTTTTTTTTGTGTTTATTGTTTATTGTTTATTGTTTCTCTCTTCCAAATCGGTAGA 1919
 QY 75 ----- 75
 Db 1920 ACAAGACCAATACACACGCGCATACTAGCCCTATTTTTTCTTGGGCTTATTATCG 1979
 QY 75 ----- 75
 Db 1980 ATTTCATTTTATTTCGAGAATATCAATGTGCGGGTTGATGTTGTTTGCATATAGTAAT 2039
 QY 75 ----- 75
 Db 2040 ACTAAACATATGCCAGTTTATACATAGATTTTTTTTAAAGATATACATGGATATGAAATG 2099
 QY 75 ----- 75
 Db 2100 AAATTTGACATTTCTCTCTTTTATTCAATATCATATATATGATCAGATACATGTACCTTT 2159
 QY 75 ----- 75
 Db 2160 TGATTTGTATATTGTTTCTTACAGTTGAAGAGAGAAATACCAAAATACCCATTTGTATA 2219
 QY 75 ----- 75
 Db 2220 TTATAGTCGGTGATGAAAGTAAATTTAACAAATTTATGATAATATAGGCCATTAATCTT 2279
 QY 76 -----AspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsn 89
 Db 2280 TGATTTTCTTTTATAGATGGGACATCAATGTTGTGATGCAATAACATCCACAC 2339
 QY 90 AspGluGlnAspGlyProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsn 109
 Db 2340 GATGAGCAAGATGGACCAAGCTTGAGAAATTTCTTGGCCGACACCAACCATTTACAAC 2399
 QY 110 ThrAsnGluAsnValGlyAspGlySerGlyCysTyrGlyGlyGlyAspGlyGly 129
 Db 2400 ACCAACGAAACGTTGGAGATGGAAAGTGGCTGTTTATGGAGGAGGAGACGGTGT 2459
 QY 130 GlyGlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProValAspAsn 149
 Db 2460 GGTGGCTCACTAGGACTTTCGATGATAAGACATGGCTGAGAAATCAACCGCTGGATAAT 2519
 QY 150 ValAspAsnGlnGluAsnGlyAsnAlaLysGlyLeuSerLeuSerMetAsnSerSer 169
 Db 2520 GTTGATAATCAAGAAATGGCAATGCTGCAAAAGGCTGTCCCTCTCAATGAATCATCT 2579
 QY 170 ThrSerCysAspAsnAsnAsnAspSerAsnAsnValValAlaGlnGlyLysThrIle 189
 Db 2580 ACTTCTTGTGATACAAACACAGACAGCAATTAACACGTTTGTGCCCAGGGAACACTATT 2639
 QY 190 AspAspSerValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSer 209

Db 2640 GATGATAGCGTTGAAGCTACACCGAAGAAACTATTGAGAGTTTTGGACAGGACGCTCT 2699
QY 210 ILeTyRArgGlyValThr----- 215
Db 2700 ATATACCGCGGTGTAC- AAGGTGCCCTTCATTATTATTAATAAATGTGTAAATGTCTG 2758
QY 215 ----- 215
Db 2759 CTTGAATTGTTATCTTCTTGGTAAAGTCTGGACATTGATCTAATGGCTCTGTTGGAGA 2818
QY 215 ----- 215
Db 2819 GTGCTACCGAATGCTCTTGATATATAGTATCAAGAGAGATATTGTTATTATGGCTTA 2878
QY 215 ----- 215
Db 2879 TATAGAATAATACATATATATATATATATACATGTTAGCTGTTGATGACATGTATGTTG 2938
QY 216 -----ArgHisArgTrrpThrGlyArgTyrGluAlaHisLeuTrpAspAsnSer 231
Db 2939 TATTAAATGATAAGGCATCGGTGGACAGGAAGATATGAGGCACATTTATGGGATAATAGT 2998
QY 232 CysLeuArgGluGlyGlnThrArgLysGlyArgGln----- 243
Db 2999 TGTAAAGAGAGAGCCCAACCGCCAAAGGAAGACAGGTATATATATATATATATATATAT 3058
QY 243 ----- 243
Db 3059 TTGATCATATTTTCATACACGATTACTTTCAAACCTAATATAGTGTTCGATCAATGTT 3118
QY 243 ----- 243
Db 3119 CATGTTTTTATCAAAATTTGCACCTGTGTGTGTTCTCAGTTTATTTGGGTAAGTAAT 3178
QY 244 -----ValTyrLeu----- 246
Db 3239 ATTTTTGTGTATCAATTCATTATATATATCAAAATAGCGATACGATCTACATTCATTTGT 3298
QY 247 -----GlyGlyTyrAspLysGluLysAlaAlaArgAlaTyrAspLeu 261
Db 3299 TGCTATATCATGCGAGAGTTATGACAAAGAGAAAGACGCTAGGCTTAATGATTTA 3358
QY 262 AlaAlaLeuLysTyrTrpGlyThrThrThrAsnPhePro----- 276
Db 3359 GCCCACTCAAGTATTGGGAAACCAACCACTACTACTAACTTCCCGTAAGTCAATCAATG 3418
QY 276 ----- 276
Db 3419 TTGTACAAGATTTTCATAACTTAGAACCAATTTTATTCTTTTATTAAGATGCTATTATC 3478
QY 276 ----- 276
Db 3479 TTATTATTAAATGCCATGTTTATATGTTTACATTTATTACAATAAAAGTACTTTTGGTT 3538
QY 277 -----MetSerGluTyrGluLysGluValGluLysMetLysHisMetThr 291
Db 3539 TGATATAATATGTAGATGAGCGAATATGAAAAAGAGGTAGAAGAGATGAAGCAGCATGACA 3598
QY 292 ArgGlnGluTyrValAlaSerLeuArgArg----- 301
Db 3599 AGGCAAGNGTATGTTGCTCACTCGCGCAGGTATATATATGNACTTCTGTATATTATTGCAT 3658
QY 301 ----- 301
Db 3659 ATGGCATCTATTATTATACATGTATATTAGTATTATATATATAGAACCCATCAGCTCAC 3718
QY 301 ----- 301
Db 3719 GTTTATATTAAAAATATGCCGTATTTCACGTGAGATTATCAGCATACACCTATATATAA 3778

QY 302 -----LysSerSerGlyPheSerArgGlyValaserIleTyrArg 314
Db 3779 TAGACATTAAATATATGAGGAAAGT-AGTGTGTTCTCTCGTGGTGCATTCATTATCGT 3837
QY 315 GlyValThr----- 317
Db 3838 GGAGTAAC- AAGGTATTTCATACAGAGAGAACGAATCTTATTTTGTAGGTACATATATAT 3896
QY 317 ----- 317
Db 3897 ATAAAAATATATATTAAGATATACATTTTATATTAATGAATATTTCTTAATGGGTCC 3956
QY 318 ---ArgHisHisGlnHisGlyArgTrrpGlnAlaArgLysGlyValAlaGlyAsnLys 336
Db 3957 AAAAGACATCACCAATCGAAGATGGCAAGCTAGGATAGGAGAGTCCCGGTAAACAA 4016
QY 337 AspLeuTyrLeuGlyThrPhe----- 343
Db 4017 GACCTCTACTTTGGGAAC-TTTTGGTACGTTTGTCTCTTACTAAACTTTCACAATCAA 4075
QY 343 ----- 343
Db 4076 ATCTATACAAAGATATCAACTAAAACTACACATATATCTAAGTAAGCTGTACATAT 4135
QY 344 -----GlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaIleLys 359
Db 4136 ATTATATATGAAGGCACACAAAGCTGCAGAGGCATACGACATTCGCGCCATCAA 4195
QY 360 PheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIle 379
Db 4196 TTCAGAGAGATTACCGCAGTGTACTTAACCTTCACATGAACAGATACACGTTTAACCAATC 4255
QY 380 LeuGluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArg 399
Db 4256 CTCGAAAGCCCTAGTCTCTTATTGTCGCGCCGCAAAACGCTCTCAAGGAGCTAACCGT 4315
QY 400 ProValProSerMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSer 419
Db 4316 CCGGTTCCAAAGPATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4375
QY 420 GlyTrpGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeuHisGln 439
Db 4376 GGTTCGCAAAACGCTGCGGTTTCAGCATCATCAGGAGTAGATTTCGAGCTTATTGACCAA 4435
QY 440 HisGlnGluArgTyrAsnGlyTyrTyrTrpAsnGlyGlyAsnLeuSerSerGluSerAla 459
Db 4436 CATCAAGAGAGTACAATGGTTATTATTACATGAGGAGAACTTCTCTTCGAGAGTGTCT 4495
QY 460 ArgAlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeu 479
Db 4496 AGGGCTGTTTCAACACAGAGGATGATCAACACCATTTCTTGAGCAACACGACAGCCTC 4555
QY 480 MetThrAsnIleAspHisGlnSerSerValSerAspAspSerValThrValCysGlyAsn 499
Db 4556 ATGACTAATATCGATCATCAAAAGTTCTGTTTCGGATGATTCGGTTACTGTTTGTGGAAT 4615
QY 500 ValValGlyTyrGlyGlyGlnGlyPheAlaAlaProValAsnCysAspAlaTyrAla 519
Db 4616 GTTGTGTTTATGTTGGTTTATCAGGATTTCAGGCCCCGGTTTAACTGCGATCCCTACGCT 4675
QY 520 AlaSerGluPheAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThr 539
Db 4676 GCTAGTGTAGTTTGAATATAACGCAAGAAACCATTTACTTTGCTCAGCAGCAGCAGACC 4735
QY 540 GlnGlnSerProGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMet 559
Db 4736 CAGCAGTCCCGCAGGTGGAGATTTTCCCGGCAATGACGAATAATGTTGGTCTTAATATG 4795
QY 560 TyrTyrHisGlyGluGlyGlyGluValAlaProThrPheThrValTrpAsnAspAsn 579
Db 4796 TATTACCATGGGAGGTGGTGGAGAGTTGCTCTCAACATTTACAGTTTGGAAACGACAT 4855

RESULT 4
 ACC00801
 ID ACC00801 standard; cDNA; 2588 BP.
 XX AC ACC00801;
 XX DT 16-MAY-2003 (first entry)
 XX XX Glycine max oil trait related cDNA sequence SEQ ID NO:351.
 DE XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW Lipo1-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CKX-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant; gene; ss.
 XX OS Glycine max.
 XX OS WO2003002751-A2.
 XX PN 09-JAN-2003.
 XX PD 27-JUN-2002; 2002WO-US020152.
 XX PF 29-JUN-2001; 2001US-0301913P.
 XX PR (DUPO) DU PONT DE NEMOURS & CO E I.
 XX PA (PION-) PIONEER HI-BRED INT INC.
 XX PI Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 PI Jones TJ, Kinney AJ, Klein TW, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;
 DR WPI; 2003-201509/19.
 DR P-PSDB; ABR40786.
 XX XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX Claim 18; Page 368; 542pp; English.
 XX CC The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, Lipo1-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity and CKX-like transcription factor activity. Also described: (1)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from antisense
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX SQ Sequence 2588 BP; 863 A; 513 C; 554 G; 658 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.19e-106 Length: 2588
 Score: 1310.00 Matches: 322
 Percent Similarity: 56.52% Conservative: 77
 Best Local Similarity: 45.61% Mismatches: 139
 Query Match: 42.00% Indels: 168
 DB: 7 Gaps: 29
 US-09-980-364-2 (1-579) x ACC00801 (1-2588)
 QY 4 AsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLysAspValTyr 23

D	b	1092	TTGAGAGGAGAGTAGTGGGTTTTCTCGCGGTGCATCCATTTATCGAGAGGTGACGAGA	1151
Q	y	319	HiShiSglnHissgIyArqTgTrpGlnAlaAargIleGlyArgValaIaGlyAsnLysAspLeu	338
D	b	1152	CACCACCAACATGGAAGGTGGCAAGCAGAGATTGGAAGAGTTGCTGCACCAACAGGATCTT	1211
Q	y	339	TyrLeuGlyThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAspIleAlaAlaIle	358
D	b	1212	TACTTGGGAACCTTTAGCACCCAAAGAGAGCGAGCGGAAGCATATGATGAGCAGCAATC	1271
Q	y	359	LysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAla	378
D	b	1272	AAATTCGAGGACTAAGTGTGTATCAAACTTTGACATGACGAGATATGACGTGAAAGC	1331
Q	y	379	IleLeuGlnSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGlu-----	396
D	b	1332	ATACTTGAGAGCACCACTTTGCCAATAGTGGTCTGCAAGCGTTTGAAGGATATGAG	1391
Q	y	397	-----AlaAsnArgProValProSerMetMet-----	405
D	b	1392	CAGTTGAACTGAGTGGATAATGGTTCATAGACAGATCAAGTAGATCATATGATCATC	1451
Q	y	406	MetIleSerAsnAsnValSerGluAsnSerAlaSerGlyTyrGlnAsnAlaAla	425
D	b	1452	ATGAGTCTCACCTTAACCAAGGAATCAATAACAACCTATGCAGGAGGGGAAACAGCAACT	1511
Q	y	426	ValGlnHishisGlnGlyValAspLeuSerLeuLeuHishisGlnGlnGluAArgTyrAsn	445
D	b	1512	-----CATCATAACTGGCAATGCTCATGTCATCCACCAACTCAACCTTGACCAACC	1565
Q	y	446	GlyTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCysPheLysGln	465
D	b	1566	ATGCATACCTTTATGACAAAGAAATTAAT-----TGTGCAAGCAAGAA	1610
Q	y	466	GluAspAspGln-----	469
D	b	1611	CAACAAGACAACTCTGATGCCCTCACTCTTTGTCTTATTCAGATATTCATCAACTTCAG	1670
Q	y	470	-----HisHisPheLeuSerAsnThrGlnSerLeu-----MetThr	481
D	b	1671	CTAGGGAACAATGGAAACACATACTCTTTCACACAAATTCAGGGTTGCACCCCTATGTTG	1730
Q	y	482	AsnIleAspHisGlnSerSerValSerAspAspSerValThrValCysGlyAsnValVal	501
D	b	1731	AGCATGGAT-----TCTGCTTCCATTGACAAATAGCTCTTCTTCTTAACTCGGTTGTT	1781
Q	y	502	-----GlyTyrGlyGlyTyrGlnGlyPhe---AlaAlaProVal---AsnCysAspAla	517
D	b	1782	TATGATGTTATGAGAGTGGTGGGGGCTACAAATGTGATGCTTATGCGCTTATGCGAACTACTACTGCT	1841
Q	y	518	TyrAlaAlaSerGluPheAspTyrAsn-----	526
D	b	1842	GTTCGTGCAAGTGATGGTGATCAAAATCCAAAGAACAAATCATGGTTTTGGTGATAATGAG	1901
Q	y	526	-----	526
D	b	1902	ATAAAGCACTGGTTATGAAAGTGTGATGGCTCTCGCACTGATCTTATCATGCAACAT	1961
Q	y	527	AlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGlnGlnSerProGlyGlyAsp	546
D	b	1962	GCAAGGAACCTGTATTATCTTACTCAACAGCAA-----	1994
Q	y	547	PheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGly	566
D	b	1995	-----TCATCTCTGTTGATACAGTGAAGGCTAGTGCAATATGATCAAGGGTCTGCATGC	2048
Q	y	567	GlyGluValAlaProThr	572
D	b	2049	AATACTTGGGTTCCAACT	2066

RESULT 5
ACCC0803
ID ACCC0803 standard; cDNA; 2463 BP.

QY 20 -----LysAspValTyrSerSerThrThrThrVal 30
Db 257 ACGGACACCTTCGTTTGGTTTCACCCCTGATGAATCTCAAGCACT- 304
QY 31 ValAspValAlaGlyGluTyrCysTyrAspProThrAlaAlaSerAspGluSerAla 50
Db 305 ---GATGTAGCAGGAGAC---TGCTTTGATCTCACTTCT- ---GACTCAACTCCT 349
QY 51 IleGlnThrSerPheProSerPheGlyValValValAspAlaPheThrArgAspAsn 70
Db 350 CATTTACTCAACCTTCCCTCT- ---TACGCG---ATATACGAGGCTTTTCATAGAGCAAC 403
QY 71 AsnSerHis---SerArgAspTrpAspIleAsn-----Gly 81
Db 404 AATATTCAACCACTCAAGATTGGAAGGAGAACTACAACACCCAAAACCTTGCTATTGGGA 463
QY 82 CysAlaCysAsnAsn-----IleHisAsn-----AspGluGlnAspGlyPro 95
Db 464 ACTTCATGCAGCAACCAAAACATGAAACCAACCAACCATCAGCAACCAACCAACACAGCCCA 523
QY 96 LysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsn---Val 114
Db 524 AAGCTTGAATACTTCTCGTGGACACTCA-----TTTGGTGAACATGAGCAACCCCTAC 577
QY 115 GlyAspGlySerGlySerGlyCysTyr-----GlyGly 125
Db 578 GGTGTAACCTCAGCCTCTACAGAATACATGTTCCCGGCTCAGCCGGTATTGGCCGGTGGC 637
QY 126 GlyAspGlyGly-----GlyGlySerLeuGlyLeuSerMet 137
Db 638 GCGCGCGTGTAGCAATAGCAGCAACACAGCAACAGTAGTCTCCATAGGGTATCCATG 697
QY 138 IleIysThrTrpLeuArgAsnGlnProValAspAsnValAspAsnGlnIleAsnGlyAsn 157
Db 698 ATAAGACATGTTGAGAACCAACCAACCACTCAGAAAACCAACAATAACAACAACAAAT 757
QY 158 AlaAlaLysGly-----LeuSerLeuSerMetAsn--- 167
Db 758 GAATGTGTGCATAGTAGAAGCAGTGTGCAGCAGACTCATCACTTTTCATAGGACT 817
QY 168 -----SerSerThr-----SerCysAspAsnAsnAsp 177
Db 818 GGTTCACAATCAAGCACATCACTACCCCTTCTCACTGCTAGTGTGGATAATAGAGAGAGT 877
QY 178 SerAsnAsnAsn-----ValValAlaGlnGlyLysThrIleAspAspSer 192
Db 878 TCTTCTGATAACAAACCAACCACTACCAACCGCTGCATTCGATAACCCCAACCGGAGCC 937
QY 193 ValGluAlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArg 212
Db 938 ATTGAAACTGCACCAGAAAGTCCATTGACACTTTTGGACAGAGAACTTCTATCTACCGT 997
QY 213 GlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCys 232
Db 998 GGTGTAAACAGGCATAGTGTGACCGGAGGTATGAGCTCACCTGTGGGATAATAGTTGT 1057
QY 233 LysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGlu 252
Db 1058 AGAAGAGAGGCAACAACTCGCAAGGAGGCAAGTTCATTGGGAGGTTATGACAAGAA 1117
QY 253 GluLysAlaAlaArgAlaTyrAspLeuAlaLeuLysTyrTrpGlyThrThrThr 272
Db 1118 GAAAGGCGAGCTAGAGCCCTAGATTGGCAGCACTAAATACTGGGGAACAACCTACGACA 1177
QY 273 ThrAsnPheProMetSerGluTyrGluLysGluValGluGluMetLysHisMetThrArg 292
Db 1178 ACAAATTTTCCAAATAGCCACTATGAGAAAGAGTGTGAAGAAATGAAGACATGACTAGG 1237
QY 293 GlnGluTyrValAlaSerLeuArgLysSerGlyPheSerArgGlyAlaSerIle 312
Db 1238 CAAGAGTACGTTGGCTCATTCAGAAAGAGAGTAGTGGGTTTTCTCGCGGGCATTCATT 1297
QY 313 TyrArgGlyValThrArgHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgVal 332

Db 1298 TATCGAGGTGTCGAGAGACACCATCAACATGGAAGATGGCAAGCGAGGATGGAAGATT 1357
QY 333 AlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaGluAla 352
Db 1358 GCTGGCAACAAGGATCTCTACTTGGAACTTTTCAGCACCCCAAGAGAGCGCAGAGAACA 1417
QY 353 TyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsn 372
Db 1418 TATGATGTAGCAATCAATTCAGAGCACTAGTGTGTGTACAACTTTTGACATGAGC 1477
QY 373 ArgTyrAsnValLysAlaIleLeuGluSerProSerLeuProIleGlySerAlaAlaLys 392
Db 1478 AGATATGACGTGAAAAGCATACTTGAGAGCACCACTTTGCCAATTTGGTGTGCTCAAG 1537
QY 393 ArgLeuLysGlu-----AlaAsnArg 399
Db 1538 CTTTTGAGGATATGAGCAGCGTGAACCTGAGGTGGAGAATGTTTCATAGACAGATCAA 1597
QY 400 ProValProSerMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSer 419
Db 1598 GAAGATCATAGTAGCATCATGAACCTCTCACTTAACCTCAAGGAATCATTAACAACCTATGCA 1657
QY 420 GlyTrpGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeuLeuHisGln 439
Db 1658 GCAGGAGGAACAACAGCGCACTCATCATCACTGACCAATGCTCTTGCAATTCACCACAA 1717
QY 440 HisGlnGluArgTyrAsnGlyTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAla 459
Db 1718 CCTCAACCTTGCAACCACTACCTTATGACAAAGAAATTAAT----- 1765
QY 460 ArgAlaCysPheLysGlnGluAspAspGln----- 469
Db 1766 -----TGTGCAAGCAAGAAACAAGCAACTCTGATGCTCTCACTTTGCTTATTCA 1819
QY 470 -----HisHisPheLeuSerAsnThrGln 477
Db 1820 GATATTCATCACTACAGTAGGGAACATGGCACACAACCTTTTTCACACAAATCA 1879
QY 478 SerLeu-----MetThrAsnIleAspHisGlnSerValSerAspAspSerValThr 495
Db 1880 GGGTGTGACCCCTATGTTAAGCATGAT-----TCTGCTTCCATTGACAATAGCTCT 1930
QY 496 ValCysGlyAsnValVal-----GlyTyrGlyGlyTyrGlnGlyPhe---AlaAlaPro 512
Db 1931 TCATCTACTCTGTTGTTTATGATGTTATGAGGTGTGGAGTGTGGGGCTAATATGATTCCT 1990
QY 513 Val---AsnCysAspAlaTyrAlaAlaSerGluPheAspTyrAsn----- 526
Db 1991 ATGGGAGCTACTACTACTGTTGTTGCAAAATGATGTGATCAAAATCCAAAGCAATCAT 2050
QY 526 ----- 526
Db 2051 GGTTTTGGTGATAATCAGATAAGCACTTGGTTATGAAGTGTGTGTTCTTACAACCT 2110
QY 527 -----AlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGln 540
Db 2111 GATCCTTATCATGCATGCAAGGAACCTGTTATTATCTTACTCAACAGCAACCATCTTCT 2170
QY 541 GlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyr 560
Db 2171 -----GTTGATGAGTGAAGGCTAGTGCATAT 2197
QY 561 TyrHisGlyGluGlyGlyGluValAlaProThr 572
Db 2198 GATCAAGGATCTGCATGCAATACCTTGGGTTCCAACT 2233
RESULT 6
ACCC0806
ID ACC00806 standard; cDNA; 1990 BP.
XX
AC ACC00806;
XX

DT 16-MAY-2003 (first entry)
 XX Glycine max oil trait related cDNA sequence SEQ ID NO:361.
 DE
 KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 XX receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW LIP15-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CKC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant; gene; ss.
 XX
 OS Glycine max.
 OS
 PN WO2003/002751-A2.
 PN
 XX 09-JAN-2003.
 PD
 XX 27-JUN-2002; 2002WO-US020152.
 PF
 XX 29-JUN-2001; 2001US-0301913P.
 PR
 XX (DUPO) DU PONT DE NEMOURS & CO E. I.
 PA (PION) PIONEER HI-BRED INT INC.
 XX
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Fanodu OO, Harvell LT;
 PI Jones TJ, Kinney AJ, Klein TW, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;
 PI
 DR WPI: 2003-201509/19.
 DR F-PSDB; ABR40771.
 DR
 XX Novel nucleotide fragment encoding polypeptides having receptor-like
 XX protein kinase activity, caleosin-like activity, useful for altering oil
 XX phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX
 XX Claim 18; Page 381-382; 542pp; English.
 XX
 CC The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, LIP15-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity, and CKC-like transcription factor activity. Also described: (1)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 1990 BP; 628 A; 391 C; 459 G; 512 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 8e-90 Length: 1990
 Score: 1119.50 Matches: 275
 Percent Similarity: 55.70% Conservative: 77
 Best Local Similarity: 43.51% Mismatches: 147
 Query Match: 35.89% Indels: 133
 DB: 7 Gaps: 23
 XX
 US-09-980-364-2 (1-579) x ACC00806 (1-1990)
 QY 1 MetAsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLys 20
 DB 126 ATGACACACACACTGGCTTTCGTTCCCTCTTCT----- 158
 QY 21 AspValTyrSerSerThrThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
 DB 158 ----- 158

Db 366 ATGAGCACATGACTAGGCAAGTAGTACGTTGGCTCATTTGAGAGGAGAGTAGTGGTTT 425
 Qy SerArgGlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHisGlyArgTrpGln 326
 Db 426 TCTCGCGGTGCATCTCTTTATCGAGAGTGACGAGACACCAACATCGAAGGTGGCAA 485
 Qy 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGln 346
 Db 486 CGAGAGATGGAAGAGTGTCTGCAACAGGATCTTTACTTGGAACTTTTAGCACCCAA 545
 Qy 347 GluGluAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaVal 366
 Db 546 GAAGAGCGAGCGAAGCATATGATGATAGCAGCAATCAAAATTCGAGGACAAAGTGTGTT 605
 Qy 367 ThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuPro 386
 Db 606 ACAAACTTTGACATGACGATATGATGAGTGAAGGATATCTTGGAGCACCACCTTTGCCA 665
 Qy 387 IleGlySerAlaAlaLysArgLeuLysGlu----- 396
 Db 666 ATAGGTGGTCTGCAAGCGCTTTGAAGGATATGAGCAGGTTGAACCTGAGTGTGGATAAT 725
 Qy 397 -----AlaAsnArgProValProSerMetMetIleSerAsnAsnValSerGlu 413
 Db 726 GGTATAGACGACATCAAGTAGATCATGATATCATCATGATTTCCACCTAACTCAAGGA 785
 Qy 414 SerGluAsnSerAlaSerGlyTyrGlnAsnAlaValGlnHisGlnGlyValAsp 433
 Db 786 ATCAATTAACAACATGATGAGAGGGGGAACAACT-----CATCATTAACGGCAACAT 839
 Qy 434 LeuSerLeuLeuHisGlnHisGlnGluArgTyrAsnGlyTyrTyrAsnGlyGlyAsn 453
 Db 840 GCTCATGCAATTCACCAACCTCAACCTTGACCAACCATGACCTTATGACCAAGA 899
 Qy 454 LeuSerSerGluSerAlaArgAlaCysPheLysGlnGluAspGln----- 469
 Db 900 ATTAAT-----TGTGCAAGCAAGCAACAACTCTGTGATGCTTCTGATGCTTCT 944
 Qy 470 -----HisHis 471
 Db 945 CACTCTTTGTCTTATTCAGATATTCATCACTTCAGTAGGGAACAATGGAACATAAC 1004
 Qy 472 PheLeuSerAsnThrGlnSerLeu-----MetThrAsnIleAspHisGlnSerSerVal 489
 Db 1005 TTCTTTCAACAATTCAGGGTTCACCCCTATGTTGAGCATGAT-----TCTGCT 1055
 Qy 490 SerAspAspSerValThrValCysGlyAsnValVal-----GlyTyrGlyGlyTyrGln 507
 Db 1056 TCATTGCAATAGCTCTCTCTTAACCTCGTGTGTTTATGATGCTTATGAGGTGTGG 1115
 Qy 508 GlyPhe---AlaAlaProVal---AsnCysAspAlaTyrAlaAlaSerGluPheAspTyr 525
 Db 1116 GGCTACAATGTGATGCTGCTATGGGAACCTACTACTGCTGTTGTTGCAAGTGTGATCAA 1175
 Qy 526 Asn----- 526
 Db 1176 ATCCAAAGAACATCATGTTTGGTGATAATGAGATAAAGCACTTGGTTATGAAGT 1235
 Qy 527 -----AlaArgAsnHisTyrTyrPheAla 534
 Db 1236 GTGTATGGCTGTGCAACTGATTTCTTATCATGCAATGCAAGGAACCTGTATTATCTTACT 1295
 Qy 535 GlnGlnGlnGlnThrGlnGlnSerProGlyGlyAspPheProAlaAlaValThrAsnAsn 554
 Db 1296 CAACGCAA-----TCATCTTCTGTGTATACA 1322
 Qy 555 ValGlySerAsnMetTyrTyrHisGlyGluGlyGlyGlyGluValAlaProThr 572
 Db 1323 GTGAAGGCTAGTGCATATGATCAAGGGTCTGCATGCAATACTTGGGTTCCAACT 1376
 RESULT 8
 ADD30878
 ID ADD30878 standard; cDNA; 1818 BP.

XX AC ADD30878;
 XX 15-JAN-2004 (first entry)
 XX Plant yield-related polynucleotide clone G1793.
 XX ds; transcription factor; transgenic plant; growth rate; senescence;
 XX seed germination rate; plant vigor; seedling vigor.
 XX Arabidopsis thaliana.
 XX WO2003013227-A2.
 XX 20-FEB-2003.
 XX 09-AUG-2002; 2002WO-US025805.
 XX 09-AUG-2001; 2001US-0310847P.
 XX 19-NOV-2001; 2001US-0336049P.
 XX 11-DEC-2001; 2001US-0338692P.
 XX 14-JUN-2002; 2002US-00171468.
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.
 XX PA Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
 XX PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pinada O, Yu G;
 XX Broun PE;
 XX WPI: 2003-248221/24.
 XX P-PSDB; ADD30879.
 XX New plant transcription factor polynucleotides and polypeptides, useful
 XX in producing transgenic plants with commercially valuable properties,
 XX PT such as an alteration in a plant growth characteristic, e.g. growth rate
 XX or apomixis.
 XX PS Disclosure; SEQ ID NO 907; 454pp; English.
 XX The invention relates to a number of isolated Arabidopsis thaliana cDNA
 XX sequences and their encoded proteins which are especially transcription
 XX factor related cDNA's and proteins. The isolated or recombinant plant
 XX transcription factor polynucleotides and polypeptides are useful in
 XX producing transgenic plants with commercially valuable properties, i.e.
 XX modified or altered desirable traits as compared to a reference plant,
 XX such as an alteration in a plant growth characteristic, e.g. growth rate,
 XX germination rate of seeds, vigor of plants and seedlings, or leaf and
 XX flower senescence. Sequence information related to the polynucleotides
 XX and polypeptides can also be used in bioinformatic search methods. The
 XX transgenic plant is useful for growing a progeny plant from a parent
 XX plant. This sequence represents one of the cDNAs of the invention.
 XX SQ Sequence 1818 BP; 551 A; 425 C; 414 G; 428 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,85e-83 Length: 1818
 Score: 1047.50 Matches: 262
 Percent Similarity: 46.9% Conservative: 70
 Best Local Similarity: 37.0% Mismatches: 110
 Query Match: 33.5% Indels: 265
 DB: 9 Gaps: 23
 US-09-980-364-2 (1-579) x ADD30878 (1-1818)
 Qy 2 AsnAsnAsnTyrLeuGlyPhe-----SerLeuSerProTyrGlu 14
 Db 65 TCTAACAACTGGCTTGGCTTTCCCTCTTACCGAACAACACTCTTCTTCCCTCTCATGAA 124
 Qy 15 GlnAsnHisHisArgLysAspValTyrSerThrThrThrValValAspValAla 34
 Db 125 TACAAAC----- 130
 Qy 35 GlyGluTyrCysTyrAspProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSer 54

```

Db 130 ----- 130
Qy 55 PheProSerProPheGlyValValValAspAlaPheThrArgAspAsnAsnSerHisSer 74
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 131 -----CTTGGCTTGGTCAGCGACCATATG-----GACACCCCTTTTCAACA 172
Qy 75 ArgAspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluInAspGly 94
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 173 CAAGAGTGGATATG-----ATCAATCCACACCGTGGAGGAGAGATCAA 217
Qy 95 -----ProLysLeuLysAsnPheLeuGlyArgThrThrThrIleTyrAsnThr 110
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 218 GGAGGAGAGGTTCCAAAGTGGCGGATTTCTCGGTGAGCAACCGGAGAAACCAA 277
Qy 111 AsnGluAsnValGlyAspGlySerGlyCysTyrGlyGlyGlyAspGlyGlyGly 130
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 278 TCCAAACCACCTAGTAGCTTACAACGACTCAGACTACTTCTCCATACCAAT----- 328
Qy 131 GlySerLeuGlyLeuSerMetIleIleIleIleIleIleIleIleIleIleIleIle 147
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 329 -----AGCTTGATGCTAGCGTCCAAATCAAACGATGCTGTGTAGCAGCT 373
Qy 148 -----AspAsnValAspAsn-----GlnGluAsnGlyAsnAla 158
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 374 TGTGACTCCATACTCTTAACAACAGTAGCTATCATGAGCTTCAAGAGAGTGCTACAAT 433
Qy 159 AlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSer 178
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 434 CTACAGTCACCTACTTTGTCTCCATGGGACCACC-----GCT 469
Qy 179 AsnAsnAsnValVal-----AlaGlnGlyLysThrIleAspAspSerValGlu--- 194
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 470 GGTATATATGTTGACAAAGCTTCACCATCCGAGACCACCGGGATACGCTAGCGGT 529
Qy 195 -----AlaThrProLysLysThrIleGluSerPheGlyGln 206
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 530 GGAGCACTAGCGGTGTTGAGACGGCCACGCAAGAGCTGATGGACACTTTCGGACAA 589
Qy 207 ArgThrSerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHis 226
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 590 CGAACCTCGATCTATCGTGGTGTACAGAGCATCGATGGAGTGGTGCATATAGAGCTCAT 649
Qy 227 LeuTrpAspAsnSerCysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeu 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 650 CTATGGGATAATAGTTGTAGAAGGGAAGGCCAGCTAGGAAGGAAGACAAAGTTTACTTG 709
Qy 247 GlyGlyTyrAspLysGluGlyLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyr 266
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 710 GGTGATATGACAAAGAGATAAAGCAGCAAGATCATATCATCTAGCTGCACCTTAAGTAC 769
Qy 267 TrpGlyThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGlu 286
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 770 TGGGGTCTTCAACTACTACTTAATTTCCCATTAACAACTACGAGAAAGATAGAGAA 829
Qy 287 MetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPhe 306
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 830 ATGAAGCAGCATGACGAGACAAAGATCTCGTGGCTGCCATTAAGAGAAAGTAGTGGATTT 889
Qy 307 SerArgGlyValaSerIleTyrArgGlyValThrArgHisGlnHisGlyValTyrGln 326
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 890 TCGAGAGCGGTTTGTATGATTCGAGGAGTTACAGGGCATCACCAACATGGAATGGCAA 949
Qy 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGln 346
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 950 GCAAGGATCGCGCGAGTCGCGGAAACAAAGACCTCTACTTGGGAACCTTTAGCACTGAG 1009
Qy 347 GluGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaVal 366
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1010 GAAGACGACGAGCAAGCTTACGATATAGTCGCAATAAAGTTTAGAGACTTAATGAGTG 1069
Qy 367 ThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuPro 386
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 1070 ACCAACTTCGAGATCAACCGGTACGACGTGAAAGCCATTTCTAGAGTAGCAGCTCTTCCC 1129
Qy 387 Ile---GlySerAlaAlaLysArgLeuLysGluAla----- 397
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1130 ATCGAGGAGGCGCGAGCTTAACCGCTCAAGAGAGCTCAAGCTCTTGAGCTTTCAGAGAAA 1189
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 397 ----- 397
Db 1190 CGCGAGGCGGAGATGATAGCCCTTGTGTTCAAGTTTCCAGTACGGTGTGGCTCGAGCACA 1249
Qy 398 -----AsnArg 399
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1250 GGCTCTGGCTCCACCTCATCAAGACTTCAGCTTCAACCTTACCCTCTAAGCATTCACAA 1309
Qy 400 ProValProSerMetMetMetIleSerAsnValSerGluSerGluSerGluSerAlaSer 419
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1310 CCATTAGAGCCTTTTCTATCTCTTCAGAACAAATGACATCTCTCATTAACAACAACAAT 1369
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 420 GlyTyrGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeu---LeuHis 438
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1370 GGTACGATTCCTCTCTTTAATCACCATAGTATATCCAGACACAACTCATCTCCAC 1429
Qy 439 GlnHis-----GlnGluArgTyrAsn 445
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1430 CAACAGACCAACAATTTACTTGCAGCAACAGTCGAGCCAGAACTCTCAGCAGCTCTACAAT 1489
Qy 446 GlyTyrTyrTyr----- 449
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1490 GGTATCTTTCATAGCAATCGGCTCTGCTTCATGGACTTGTTCTTACCTCTATCGTTGAC 1549
Qy 450 -----AsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCysPheLysGln 465
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1550 AACAATATAACAATGAGGCTTAGTGGAGCTACACACTGCAGCA----- 1597
Qy 466 GluAspAspGlnHisPheLeuSerAsn----- 475
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1598 -----TTTCTTGGAAACCAACCGGTATTGGTATTGGTCCAGCTCGACT 1639
Qy 476 -----ThrGlnSerLeuMetThr---AsnIleAspHisGlnSerSerValSerAsp 491
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1640 GTTGATCGACGAGAGTTTCCAAACCGTTAAACAGATTAGCATATGCTTCCAGTGAT 1699
Qy 492 AspSerValThrValCysGlyAsnValValGlyTyrGlyGlyGlnGlyPheAlaAla 511
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1700 -----GGAACCGGAGGTATAGTGGTTGG----- 1723
Qy 512 ProValAsnCysAspAlaTyrAlaAlaSerGluPheAspTyrAsnAlaArgAsnHisTyr 531
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1723 ----- 1723
Qy 532 TyrPheAlaGlnGlnGlnThrGlnGlnSerProGlyGlyAspPheProAlaAlaMet 551
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1724 -----ACCACTGAGTCTGTTACAGGGGTCAACACCT----- 1753
Qy 552 ThrAsnAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGlyGluValAlaPro 571
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1754 -----GGTGGT----- 1759
Qy 572 ThrPheThrValTrpAsnAsp 578
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1760 GTTTTCACTATGTGAATGAG 1780

```

RESULT 9

ADD30822
ID ADD30822 standard; cDNA; 1818 BP.

XX AC ADD30822;

XX DT 15-JAN-2004 (first entry)

XX DE Plant yield-related polynucleotide clone G1793.

XX KW ds; transcription factor; transgenic plant; growth rate; senescence;

```

KW seed germination rate; plant vigor; seedling vigor.
XX Arabidopsis thaliana.
XX WO2003013227-A2.
XX 20-FEB-2003.
XX 09-AUG-2002; 2002WO-US025805.
XX 09-AUG-2001; 2001US-0310847P.
XX 19-NOV-2001; 2001US-0336049P.
XX 11-DEC-2001; 2001US-0338692P.
XX 14-JUN-2002; 2002US-00171468.
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
XX Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;
XX Brown PE;
XX WPI; 2003-248221/24.
XX P-PSDB; ADD30823.
XX New plant transcription factor polynucleotides and polypeptides, useful
XX in producing transgenic plants with commercially valuable properties,
XX such as an alteration in a plant growth characteristic, e.g. growth rate
XX or apomixis.
XX Disclosure; SEQ ID NO 851; 454bp; English.
XX The invention relates to a number of isolated Arabidopsis thaliana cDNA
XX sequences and their encoded proteins which are especially transcription
XX factor related cDNA's and proteins. The isolated or recombinant plant
XX transcription factor polynucleotides and polypeptides are useful in
XX producing transgenic plants with commercially valuable properties, i.e.
XX modified or altered desirable traits as compared to a reference plant,
XX such as an alteration in a plant growth characteristic, e.g. growth rate,
XX germination rate of seeds, vigor of plants and seedlings, or leaf and
XX flower senescence. Sequence information related to the polynucleotides
XX and polypeptides can also be used in bioinformatic search methods. The
XX transgenic plant is useful for growing a progeny plant from a parent
XX plant. This sequence represents one of the cDNAs of the invention.
XX SQ Sequence 1818 BP; 551 A; 425 C; 414 G; 428 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.85e-83 Length: 1818
Score: 1047.50 Matches: 262
Percent Similarity: 46.96% Conservative: 70
Best Local Similarity: 37.06% Mismatches: 110
Query Match: 33.58% Indels: 265
DB: 9 Gaps: 23

US-09-980-364-2 (1-579) x ADD30822 (1-1818)
Qy 2 AsnAsnAsnTrpLeuGlyPhe-----SerLeuSerProTyrGlu 14
Db 65 TCTAACAACTGGCTTGGCTTTCCTTTCCACGAAACAACCTCTCTTTGGCTCTCCATGAA 124
Qy 15 GlnAsnHisHisArgLysAspValTy-SerSerThrThrThrThrValValAspValAla 34
Db 125 TACAAAC----- 130
Qy 35 GlyGlutyrCysTyrAspProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSer 54
Db 130 ----- 130
Qy 55 PheProSerProPheGlyValValValAspAlaPheThrArgAspAsnAsnSerHisSer 74
Db 131 -----CTTGCTTGGTCAGCGCCCATATG-----GACAACCCCTTTTCAAAACA 172
Qy 75 ArgAspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGly 94

```


Db 1645 -----CCTCAAGCCCCAGTGTGGAGTCAGTGGACG 1679
QY 523 eAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGlnGlnSerPr 543
Db 1680 AGATTAT-----GGACATGGCTACTTCCACCTTGATGGACCAAAATATGATGATG 1730
QY 543 oGlyGlyAspPheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyrTyrHisG1 563
Db 1731 GACAATGAAATGACACATGACGACCAACAGCTTGGAAAT----- 1773
QY 563 yGluGlyGlyGlyGluValAlaProThrPheThrValTrpAsnAsp 578
Db 1774 -TTGGATTGGTTAATCAAGTTCCTCATGTTTGTCTTGTGGATGAA 1818
RESULT 15
ID ACC00799 standard; cDNA; 1874 BP.
XX AC ACC00799;
XX DT 16-MAY-2003 (first entry)
XX DE Glycine max oil trait related cDNA sequence SEQ ID NO:347.
XX KW plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
KW receptor-like protein kinase; mitogen activated protein kinase; oil;
KW Lip15-like transcription factor caleosin; ATP citrate lyase; SNF1;
KW CKC-like transcription factor; antisense inhibition; co-suppression;
KW transgenic plant; gene; ss.
XX OS Glycine max.
XX PN WO2003002751-A2.
XX PD 09-JAN-2003.
XX PF 27-JUN-2002; 2002WO-US020152.
XX PR 29-JUN-2001; 2001US-0301913P.
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Allen SM, Allen WB, Epelbaum S, Famodu OO, Harvell LT;
PI Jones TJ, Kinney AU, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
PI Tarczynski MC;
XX WP: 2003-201509/19.
XX DR P-PSDB; ABR40764.
XX PT Novel nucleotide fragment encoding polypeptides having receptor-like
PT protein kinase activity, caleosin-like activity, useful for altering oil
PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
XX PS Claim 18; Page 364-365; 542pp; English.
XX CC The present invention describes an isolated nucleotide fragment (I)
CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
CC activated protein (MAP)-kinase activity, Lip15-like transcription factor
CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
CC activity and CKC-like transcription factor activity. Also described: (1)
CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
CC oil obtained from (V). (I) or its part can be used in antisense
CC inhibition or co-suppression in a transformed plant. (III) is useful for
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
CC creating transgenic plants having altered lipid profiles. (I) can also be
CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
CC ABR40879 represent sequences used in the exemplification of the present

CC invention
XX SQ Sequence 1874 BP; 542 A; 472 C; 412 G; 448 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 4.28e-78 Length: 1874
Score: 987.50 Matches: 234
Percent Similarity: 61.64% Conservative: 60
Best Local Similarity: 49.06% Mismatches: 107
Query Match: 31.66% Indels: 78
DB: 7 Gaps: 13
US-09-980-364-2 (1-579) x ACC00799 (1-1874)
QY 4 AsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisHisArgLysAspValTyr 23
Db 184 AACTGGCTTTGGTCTCTCTCTCTCCCAATGGAATGCTCCGAACCTCCGACCTCAGTTC 243
QY 24 SerSerThrThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp---ProThr 42
Db 244 CTTCAA-----TACGACGCCGCTTCC 264
QY 43 AlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGlyValVal 62
Db 265 GCTACTTCTCATCATCACTACTACTCTGCACACTGTGTACACCAACGGGTGGGC----- 318
QY 63 ValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAspIleAsnGlyCys 82
Db 319 -----AACGGAGCCTCAAGTTTGACGAGATCTCAACACAGT 357
QY 83 AlaCysAsnAsnIleHisAsnAspGluGlnAsp-----GlyPro 95
Db 358 GACGTGAGTTTGTGAATCTTGTGCGCAGAGCGTGGCCACGTCGCCCGCCGCGCGCG 417
QY 96 LysLeuGluAsnPheLeuGlyArgThr----- 104
Db 418 AAGCTGGAGGATTTCTCGGCGACTCTCCGCGGTGATCGGTACTCCGACAGCCAGACG 477
QY 105 -----ThrThrIleTyrAsnThrAsnGluAsnValGlyAspGly 117
Db 478 GAGACGCGAGGACTCGCTCGACGACATCTACGACACCCACCACCCACCCACCCAC 537
QY 118 SerGlySerGlyCysTyrGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMet 137
Db 538 CACGGTTCTACTTCGTACTTCGTGCTGACCA-----GCAGGATCTCAAGGCCATTACTGG 593
QY 138 IleLysThrTrpLeuArgAsnGlnProValAspAsnVal-----AspAsnGln 153
Db 594 ATTCAAGCTTTT-CGACCAACTCCGTTCCGAGGTTGATGATTCGATCCATCCATCGGAA 652
QY 154 GluAsnGlyAsnAlaAlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCysasp 173
Db 653 AAGCGCGCCGACGAGTTCGGGACTCTACTATTGAGTCTCCCGCAACGAGTTCCGCG 712
QY 174 AsnAsnAsnAspSer-AsnAsnAsnValValAlaGlnGlyLysThr---Ile-AspAsps 192
Db 713 CGTTCTCCGGTGGCACACCGGAACCTTGTCTCGCTCGCGTTGCTGACTGAGCTCCGAGAAG 772
QY 192 erValGluAla-----ThrProLysLysThrIleGluSerPheGlyGlnArgT 208
Db 773 CGGTGTCGCGCGGAGTCCAAATAGCTCGAAGAGATCTGGATACCTTCGCCCGCAGCGA 832
QY 208 hrSerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHisLeuT 228
Db 833 CTTCTATTACAGAGGTGTTACTAGGCACCGATGACAGGAAGATATGAAGCGCATCTAT 892
QY 228 rAspAsnSerCysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyG 248
Db 893 GGGACAATAGTTCGACAGGGAGGCTCAAGCAGAAAAGGGCGTCAAGTTTATTGGGGTG 952
QY 248 lYTrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpG 268
Db 953 GATATGATAGAGAAAGAAAGGCGCGGAGAGCTTATGATTGGCAGCTCTAAAGTACTGG 1012

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 19, 2004, 15:02:27 ; Search time 124 Seconds
(without alignments)
2591.263 Million cell updates/sec

Title: US-09-980-364-2

Perfect score: 3119

Sequence: 1 MNXNLGFLSLPYEQNHHRK.....YHGGGGGVAFTFTVNDN 579

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlh
-Q/cgn2_1/USFTO.spool/US0980364/runat 15072004.095237 25896/app query.fasta_1.775
-DB=issued_Patents.NA -QFMT=fastap -SUFFIX=rni -XINMATCH=0.1 -LOEFL=0
-LOEFL=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human0.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0980364.cgn 1.1 69 @runat 15072004.095237 25896 -NCPU=6
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/6C COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	913.5	29.3	1738	4	US-09-227-421-4
2	913.5	29.3	1738	4	US-09-479-855-4
3	891	28.6	2148	4	US-09-227-421-1
4	891	28.6	2148	4	US-09-479-855-1
5	461.5	14.8	1680	2	US-08-700-152A-3
6	262	8.4	301	4	US-09-313-294A-6727
7	240.5	7.7	11721	4	US-09-026-039-3
8	240.5	7.7	11721	4	US-09-026-039-3
9	179	5.7	294	4	US-09-313-294A-6087
10	155	5.0	751	4	US-09-533-029-81
11	154.5	5.0	1265	4	US-09-964-850-3
12	152.5	4.9	1577	4	US-09-533-029-11
					Sequence 4, Appli
					Sequence 1, Appli
					Sequence 3, Appli
					Sequence 6727, Ap
					Sequence 3, Appli
					Sequence 6087, Ap
					Sequence 81, Appli
					Sequence 11, Appli

13	150	4.8	1055	4	US-09-533-029-21	Sequence 21, Appli
14	149.5	4.8	987	4	US-09-964-850-5	Sequence 5, Appli
15	145	4.6	1405	4	US-09-202-161B-6	Sequence 6, Appli
16	144.5	4.6	921	4	US-09-964-850-1	Sequence 1, Appli
17	141.5	4.5	1132	3	US-08-894-731-3	Sequence 3, Appli
18	140	4.5	1669	4	US-09-026-039-1	Sequence 1, Appli
19	138.5	4.4	806	4	US-09-198-119C-78	Sequence 78, Appli
20	135	4.3	651	4	US-09-198-119C-14	Sequence 14, Appli
21	135	4.3	933	4	US-09-301-666A-1	Sequence 1, Appli
22	135	4.3	933	4	US-09-301-217-1	Sequence 1, Appli
23	134.5	4.3	563	4	US-09-198-119C-64	Sequence 64, Appli
24	132.5	4.2	1281	4	US-09-533-029-65	Sequence 65, Appli
25	132	4.2	913	4	US-09-533-029-33	Sequence 33, Appli
26	131	4.2	577	4	US-09-198-119C-38	Sequence 38, Appli
27	129.5	4.2	1500	3	US-09-300-672-1	Sequence 1, Appli
28	128.5	4.1	174	4	US-09-202-161B-21	Sequence 21, Appli
29	127.5	4.1	761	4	US-09-202-161B-5	Sequence 5, Appli
30	126.5	4.1	651	4	US-09-198-119C-12	Sequence 12, Appli
31	126.5	4.1	944	4	US-09-301-666A-7	Sequence 7, Appli
32	126.5	4.1	944	4	US-09-301-217-7	Sequence 7, Appli
33	125	4.0	349	4	US-09-198-119C-94	Sequence 94, Appli
34	125	4.0	877	4	US-09-198-119C-72	Sequence 72, Appli
35	125	4.0	2161	4	US-09-336-115C-17	Sequence 17, Appli
36	124.5	4.0	884	4	US-09-198-119C-76	Sequence 76, Appli
37	124.5	4.0	2270	4	US-09-336-115C-13	Sequence 13, Appli
38	124	4.0	905	2	US-08-949-603-1	Sequence 1, Appli
39	124	4.0	905	2	US-08-706-270A-1	Sequence 1, Appli
40	124	4.0	905	2	US-08-949-580-1	Sequence 1, Appli
41	124	4.0	905	2	US-08-950-172A-1	Sequence 1, Appli
42	124	4.0	905	4	US-09-198-119C-1	Sequence 1, Appli
43	124	4.0	937	4	US-09-301-666A-5	Sequence 5, Appli
44	124	4.0	937	4	US-09-301-217-5	Sequence 5, Appli
45	124	4.0	964	4	US-09-533-029-17	Sequence 17, Appli

ALIGNMENTS

RESULT 1
US-09-227-421-4
; Sequence 4, Application US/09227421
; Patent No. 6559357
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Mizukami, Yukiko
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility
; FILE OF INVENTION: and Enhancing Asexual Reproduction in Plants
; FILE REFERENCE: 023070-090700PC
; CURRENT APPLICATION NUMBER: US/09/227,421
; CURRENT FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: US 09/227,421
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1738
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1647)
; OTHER INFORMATION: BANT, ANT ortholog from Brassica napus (Canola)
US-09-227-421-4

Alignment Scores:
Pred. No.: 5.7e-88
Score: 913.50
Percent Similarity: 53.00%
Best Local Similarity: 41.97%
Query Match: 29.2%
Indels: 131
Gaps: 14

US-09-980-364-2 (1-579) x US-09-227-421-4 (1-1738)

Qy	275	PhaProMetSerGluTyrGluLysGluValGluGluMetThrArgGlnGlu	294
Db	1018	TTCTCTGTGGAGAAATTATCATGAAGAAGAGATTGATGACATGAAGAACAATGACTCGACCAAGAA	1077
Qy	295	TyrValAlaSerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArg	314
Db	1078	TATGTTGCTCATTGAGAGAAAACACAGTGGTTTCTTAGGGGTCTTCCATCTATAGA	1137
Qy	315	GlyValThrArgHisGlnHisGlyArgTrrGlnAlaAargIleGlyArgValAlaGly	334
Db	1138	GGAGTCACCAGACATCACCAGCATGGAAGGTGGCAAGCTCGGATCGGTAGAGTCGCTGGA	1197
Qy	335	AsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAsp	354
Db	1198	AACAAGAATCTCTACCTTGGAACTTTCGAACTCAAGAGAGAGCGCGGAGACCTATGAT	1257
Qy	355	IleAlaAlaIleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyr	374
Db	1258	GTACGACGCTATCAAGTTCGCTGGCACAAACGGCGTGACTTAACCTTGCACATACAAGGTAC	1317
Qy	375	AsnValLysAlaIleLeuGluSerProSerLeuProIleGlySerAlaAlaLysArgLeu	394
Db	1318	GATGTTGATCGCATAAGCTAGTAACACTCTCTTCTTGGAGAGATGGCTCAAGG---	1374
Qy	395	LysGluAlaAsnArgProValProSerMetMetMetIleSerAsnAsnValSerGluSer	414
Db	1375	-----AACAGCAACAGCATCGTGGTCCGCAACATTAGCGACGAG	1413
Qy	415	GluAsnSerAlaSerGlyTrpGlnAsnAlaAlaValGlnHisGlnGly	431
Db	1414	GAAGCCGCTTTAAACCCCTCTGTGAACGGTGGTTCATTAAGGAAGTGGGT	1464

RESULT 2

US-09-479-855-4

; Sequence 4, Application US/09479855

; Patent No. 6639128

; GENERAL INFORMATION:

; APPLICANT: Fischer, Robert L.

; APPLICANT: Mizukami, Yukiko

; APPLICANT: The Regents of the University of California

; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility

; TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants

; FILE REFERENCE: 023070-090720US

; CURRENT APPLICATION NUMBER: US/09/479,855

; CURRENT FILING DATE: 2000-01-07

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 4

; LENGTH: 1738

; TYPE: DNA

; ORGANISM: Brassica napus

; FEATURE:

; OTHER INFORMATION: canola AINTEGUMENTA (ANT) partial cDNA including

; OTHER INFORMATION: coding region

; NAME/KEY: CDS

; LOCATION: (1)..(1647)

; OTHER INFORMATION: canola AINTEGUMENTA (ANT)

US-09-479-855-4

Alignment Scores:		
Pred. No.:	5,7e-88	Length: 1738
Score:	913.50	Matches: 217
Percent Similarity:	53.00%	Conservative: 57
Best Local Similarity:	41.97%	Mismatches: 112
Query Match:	29.29%	Indels: 131
DB:	4	Gaps: 14

US-09-980-364-2 (1-579) x US-09-479-855-4 (1-1738)

Qy	4	AsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArg	19
Db	40	AATTTCGTAGGTTCTCGTTGCTT-----TCAAATATCTTGAATAATGGTGGTGA	90

RESULT 7

```
US-09-026-039-3
; Sequence 3, Application US/09026039
; Patent No. 6329567
; GENERAL INVENTOR:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamoto, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,039
; FILING DATE: 19-FEB-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,272
; FILING DATE: 15-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/879,827
; FILING DATE: 20-JUN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/700,152
; FILING DATE: 20-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11721 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..11721
; OTHER INFORMATION: /note= "Arabidopsis APETALA2 (AP2)"
; OTHER INFORMATION: Genomic sequence"
US-09-026-039-3

Alignment Scores:
Pred. No.: 1.56e-14 Length: 11721
Score: 240.50 Matches: 129
Percent Similarity: 32.29% Conservative: 57
Best Local Similarity: 22.40% Mismatches: 115
Query Match: 7.71% Indels: 276
DB: 4 Gaps: 19

US-09-980-364-2 (1-579) x US-09-026-039-3 (1-11721)

Qy 7 GlyPheSer-----LeuSerProTyrGluGlnAsnHis----- 17
Db 995 GGTTCCTCTATTAGCTTCTAACCTTGAGAGAAATCATACCAGAGGATTGAAGTTTGA 1054
Qy 18 -----HisArgLysAspValTyrSerSerThr 26
Db 1055 CTTTCAAAGATCAAAATCAAGAAACCAAAAAAATGTGGGATCTAAACGA 1114
```

```
Qy 27 ThrThrThrValValAspValAlaGlyGluTyrCysTyrAspPro-ThrAlaAlaSerAs 46
Db 1115 CGCACCA-----CCAAACACAAAGAGAGA 1141
Qy 46 pGluSerSerAlaIleGlnThrSerPheProSerProPheGlyValValValAspAlaPh 66
Db 1142 AGAATCTGAAGAGTTTCTTATCTTCCACCAAGTAA-----CGGGTTGGATCTTT 1192
Qy 56 eThrArgAspAsnAsnSer-----HisSerArgAspTrpAspIleAs 80
Db 1193 CTCTAATCTAGCTCTCTAGCTGTTGTATCGAAGATGGATCCGATGACGATGAACCTAA 1252
Qy 80 nGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsnPh 100
Db 1253 CCGGTCAGACCCCAATAACCCACTGTGCACCATCAGTCTCTCCCTGAGATGGATTCT-- 1310
Qy 100 eLeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGlySe 120
Db 1310 ----- 1310
Qy 120 rGlyCysTyrGlyGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMetIleLysTh 140
Db 1311 -----AACGCGGTGGTGTCTTCTGGCTTCTCTCGGGCTCAC-- 1349
Qy 140 rTrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaL 160
Db 1350 -TGGTTT----- 1355
Qy 160 sGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAsnAspSerAsnAs 180
Db 1356 -GGTGTAAAGTTTGTCTAGTCGGATCTAGCCACCGGATCGTCCGGGGTAAAGCTACCA 1414
Qy 180 nAsnValValAlaGlnGlyLysThrIleAspAspSerValGlu---AlaThrPro---Ly 198
Db 1415 CGTTCGCCGCTGCC-----GTAGTGGAGCGGCACACCGCTTGA 1453
Qy 198 sLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisAr 218
Db 1454 AAAGAGTCGGCGTGACCAAGATCAAGAGTCTCAGTATAGAGGTGTACGTTTACCG 1513
Qy 218 qTrpThrGlyArgTyrGluAlaHisLeuTrp-AspAsn----- 230
Db 1514 CGGTACCGAAGATGGGAATCTCATATTTGGTAATAATCTCATATTTTAAATTCGTAA 1573
Qy 231 -----SerCysLysArgGluGln- 237
Db 1574 TCGATCGTACTTTAGATTATAAATTTAAGTTTTTTTTTTTGTGTCTCTCTGAATTCAG 1633
Qy 238 -----ThrArgLysGlyArgGlnValTyr----- 245
Db 1634 GACTGTGGAAACAAAGTTTACTTAGTAAATTTTATTTTCTCATGTTTTTTTTTGTATT 1693
Qy 246 -----LeuGlyGlyTyrA 250
Db 1694 TGGTGTGAAAAATGTCATCATATAATTTTAATTATTATTAACTCTGAATAGTGGATTG 1753
Qy 250 sPlysGluGluLysAlaAla----- 256
Db 1754 ACATGCTCATGCGACGACG-TCGGGTATTTTCTCTCTTTGACTCTCTCTATATTGAGTTG 1812
Qy 256 ----- 256
Db 1813 TTATTTATTTATTTTAAAAATACCGAAGAAATTTATAAAATTAATTTAATTTT 1872
Qy 257 -----ArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyThrThrThrThr 273
Db 1873 TTTTATTTAATAGACATATGATAGCTGTCTATTAATTCCTGGGATAGAACGGATA 1932
Qy 273 hrAsnPheProMetSerGluTyrGluLysGluValGluGlu----- 286
Db 1933 TCAATTTCAATCGAAGATTATGATGATGACTTGAACACAGGTAAATAATAATTATAAC 1992
```

QY 286 ----- 286
Db 1993 TATATTGTTTTTATTACGATTTTAAAGGTTTGGGAGATTATATTGAAATTGAATTT 2052
QY 287 -----MetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgLysSerSerg 305
Db 2053 TATAGATGACTAAATTTACCAAGGAAGAGTTTCGTACAGTACTTCGCCGCAAGAAGCACAG 2112
QY 305 LYPheSerArgGlyValaSerIleTyrArgGlyValThrArgHisHisGlnHisGlyValArg 325
Db 2113 GCTTCCCTCGAGGAAGTTCGAAGATATAGAGGTGTCACCTTTG---CATAGTGTGGTCTT 2169
QY 325 TPGLnAlaArgIleGlyArgValAlaGly----- 334
Db 2170 GGAAGCTCGAATGGTCAATTTCTAGCAAAAGTATAATTTCTCTCATTTTATATATCA 2229
QY 334 ----- 334
Db 2230 CTCGAAAACTTCATTTTAGTTTGTATTTAACTTTGAGTTTGTGTTTCTTGAACTTA 2289
QY 335 ---AsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaAla----- 350
Db 2290 TAAAAAGTAGTATGTTTATTTGGTGTGTTGTCACACCGAGGTGGAAGCTGCTAGGTAATG 2349
QY 350 ----- 350
Db 2350 TCTTTTGTGTTGATCTCTACAAACACACATGTTGTATATGTTTGTGTTTCTCGTACTAATT 2409
QY 351 -----GluAlaTyrAspIleAlaAlaIleLysPheArg 361
Db 2410 GATTTTCATTTTATATATATATACACAGAGCTTACGATAAGCTGCAATCAATGTAAC 2469
QY 362 GlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsn 375
Db 2470 GCGAAAGACGCGGTGACCAACTTTGATCCGAGTATTTACGAT 2511

RESULT 8

US-09-026-039-3/c
; Sequence 3, Application US/09026039
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,039
; FILING DATE: 19-FEB-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,272
; FILING DATE: 15-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/879,827
; FILING DATE: 20-JUN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/700,152
; FILING DATE: 20-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774

REFERENCE/DOCKET NUMBER: 023070-067230US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 11721 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: -
LOCATION: 1..11721
OTHER INFORMATION: /note= "Arabidopsis APETALA2 (AP2)
OTHER INFORMATION: genomic sequence"
US-09-026-039-3
Alignment Scores:
Pred. No.: 1,56e-14 Length: 11721
Score: 240.50 Matches: 129
Percent Similarity: 32.23% Conservative: 57
Best Local Similarity: 22.40% Mismatches: 115
Query Match: 7.71% Indels: 276
DB: Gaps: 19
US-09-980-364-2 (1-579) x US-09-026-039-3 (1-11721)
QY 7 GlyPheSer-----LeuSerProTyrGluInAsnHis----- 17
Db 10727 GGTTCCTCTCTATTAGTCTTAACCTTCAGGAGATCATCCAGAGGATTGAAGTTTGAA 10668
QY 18 -----HisArgLysAspValTyrSerSerThr 26
Db 10667 CTTCAAAGATCAAATCAAGAAACCAAAAAAATGTCGGATCTAAACGA 10608
QY 27 ThrThrValValAspValAlaGlyGluTyrCysTyrAspPro-ThrAlaAlaSerAs 46
Db 10607 CGCACCACA-----CCAAACACAAAGAGAAGA 10581
QY 46 pGluSerSerAlaIleGlnThrSerPheProSerProPheGlyValValValAlaAlaPh 66
Db 10580 AGAATCTGAAGAGTTTGTGTTTCTTCCACCAAGTAAA-----CGGTTGATCTTT 10530
QY 66 eThrArgAspAsnAsnSer-----HisSerArgAspTrpAspIleAs 80
Db 10529 CTCTAATCTAGCTCTTTCAGCTGTTGTTATCGAAGATGATCCGATGACGATTA 10470
QY 80 nGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsnPh 100
Db 10469 CCGGGTCAGACCCCAATAACCCCACTTGTCCACCATCAGTTCTTCCCTGAGATGATCTT 10412
QY 100 eLeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGlySe 120
Db 10412 ----- 10412
QY 120 rGlyCysTyrGlyGlyAspGlyGlyGlySerLeuGlySerLeuGlyLeuSerMetIleLysTh 140
Db 10411 -----AACGGCGGTGGTGTGCTTCTGCTTCCTCGGGCTCAC-- 10373
QY 140 rTrpLeuArgAsnGlnProValAspAsnValAspAsnGlnGluAsnGlyAsnAlaAlaLy 160
Db 10372 -TGGTTT----- 10367
QY 160 sGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAsnAspSerAsnAs 180
Db 10366 -GGTGTAAAGTTTTCGTAGTCGGATCTAGCCACCGGATCGTCGCGGGAAGACTACCAA 10308
QY 180 nAsnValValAlaGlnGlyLysThrIleAspAspSerValGlu---AlaThrPro---Ly 198
Db 10307 CGTTCGCCGTGCC-----GTAGTGGAGCCGCGCACAGCGGTGAA 10269
QY 198 sLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyValThrArgHisAr 218

Db 10268 AAGAGTCGGCGTGGACCAAGATCTCAGATAGAGGTGTTACGTTTACCG 10209
Qy 218 gTTPThrGlyArgTyGluAlaHisLeuTrp AspAsn----- 230
Db 10208 GCGTACCGGAAGATGGGAATCTCATATTTGGTAATAATCTCPATTTTAAATTCGTTAA 10149
Qy 231 -----SerCysLysArgGluGlyGln- 237
Db 10148 TCGATCGTACTTTAGATTATAAATTTAAAGTTTTTTTGTGTTCTCTCGAATTCAGG 10089
Qy 238 -----ThrArgLysGlyArgGlnValTyr----- 245
Db 10088 GACTGTGGGAACAAGTTTACTTTAGTAAATTTTATTTTCTCATGTTTTTTTGTATTT 10029
Qy 246 -----LeuGlyGlyTyrA 250
Db 10028 TGTGTGTGAAATATGCATCATATAATTTTAAATTTTATTAACTCTCTGAATGGTAAATTTG 9969
Qy 250 sPLySGluGluLysAlaAla----- 256
Db 9968 ACACGTCTCATGCGAGCAGC-TCGGTATTTTCTCTCTCTGACTCTCTCTATATTGAGTTG 9910
Qy 256 ----- 256
Db 9909 TTATTTATTTATTTTAAATAATACCGGAAGAAATTTTAAATAATTTTAAATTTG 9850
Qy 257 -----ArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyThrThrThrT 273
Db 9849 TTTTATTTATAGAGCATATGATAGAGTCTCTATTAATTTCCGTGGAGTAGAAGCGGATA 9790
Qy 273 hrAsnPheProMetSerGluTyGluLysGluValGluGlu----- 286
Db 9789 TCAATTTCAACATCGAAGATTATGATGATGACTTGAACAGAGTAAATATATAAATTTATAAAC 9730
Qy 286 ----- 286
Db 9729 TATATCGTGTTTTATTAACGATTTTAAAGGTTTGGAGATTAAATTTGAATTTGATTT 9670
Qy 287 -----MetLysHisMetArgGlnGluTyrValAlaSerLeuArgLysSerSerG 305
Db 9669 TATAGATGACTAAATTTAAACAAAGAGAGTTCGTACACGACTTCGCCGCAAAAGCACAG 9610
Qy 305 LysPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHisGlyArgT 325
Db 9609 GCTTCCTCGAGGAAGTTTGAAGTATAGAGGTGTCACCTTTG---CATAAGTGTGTCGTT 9553
Qy 325 rpGlnAlaArgIleGlyArgValAlaGly----- 334
Db 9552 GGGAGCTCGAATGGGTCAATTTCTAGGCAAAAGTAAATTTCTCTCATTTTATTTCA 9493
Qy 334 ----- 334
Db 9492 CTCGAAACACTTATTTTGTATTTTAACTTTTGTGTTTGTCTTGTGAATCTTA 9433
Qy 335 ---AsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluAlaAla----- 350
Db 9432 TAAATAGGTATGTTTATTTGGTTTTCGACACCGAGTCGAAAGCTGTAGTAAATG 9373
Qy 350 ----- 350
Db 9372 TCITTTTGTGTTGATTCTACAACACACATTTGTTGTAATGTTTCTCGTTACTAAT 9313
Qy 351 -----GluAlaTyrAspIleAlaAlaIleLysPheArg 361
Db 9312 GATTTTCATTTTATATATATATACACAGAGCTTACGATAAGCTGCAATCAAAATGTAAC 9253
Qy 362 GlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsn 375
Db 9252 GGCAAGACGCGGTGACCAACTTTGATCCGAGTATTACGAT 9211
RESULT 9
US-09-313-294A-6087

; Sequence 6087, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Laligudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6087
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6476212 700351215H1
US-09-313-294A-6087

Alignment Scores:
Pred. No.: 1,94e-10 Length: 294
Score: 179.00 Matches: 42
Percent Similarity: 56.47% Conservative: 6
Best Local Similarity: 49.41% Mismatches: 4
Query Match: 5.74% Indels: 33
DB: Gaps: 1

US-09-980-364-2 (1-579) x US-09-313-294A-6087 (1-294)
Qy 251 LysGluGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyThrThr 270
Db 5 AAAGAAGAGAAGCTTACCGGGCTTACAATTTAGCTGCTCTCAAGTACTGGGAACACCA 64
Qy 271 ThrThrThrAsnPhePro----- 276
Db 65 AGCACTACAGTTTCCAGTATTTGCATATATTTTCAGTCTCTCAAGATATATCTT 124
Qy 277 -----MetSe 278
Db 125 TTTTTCCTTTTACACTGCTGCTTTGGATATATATACGATTTATTTTTCATGCTAGATCAG 184
Qy 278 rGluTyrGluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlase 298
Db 185 CAACTATGAGAAGGACCTTGGAGAGATGAATATATATGTGACGACAAAGATATGTTGCGTC 244
Qy 298 rLeuArgArgLys 302
Db 245 ACTTAGAAGGAAG 257

RESULT 10
US-09-533-029-81
; Sequence 81, Application US/095333029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddle, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22

; EARLIER APPLICATION NUMBER: 60/125,814
 ; EARLIER FILING DATE: 1999-03-23
 ; NUMBER OF SEQ ID NOS: 121
 ; SOFTWARE: PatentIn version 2.1
 ; SEQ ID NO 81
 ; LENGTH: 751
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: G25
 ; LOCATION: (75)..
 ; OTHER INFORMATION: G25
 US-09-533-029-81

Alignment Scores:
 Pred. No.: 3,136-07 Length: 751
 Score: 155.00 Matches: 50
 Percent Similarity: 45.08% Conservative: 37
 Best Local Similarity: 25.91% Mismatches: 94
 Query Match: 4.97% Indels: 12
 DB: 4 Gaps: 4

US-09-980-364-2 (1-579) x US-09-533-029-81 (1-751)

QY 265 LysTyrTrpGlyThrThrThrThrAsnPheProMetSerGluTyrGluLysGluVal 284
 DB 77 AAAATGTGGGGAGCTATCATTTCTGATTCATCTGGTGAATCTGAGTCAGAACCG 136
 QY 285 GluGluMetLysHisMetThr-----ArgGlnGluTyrValAlaSerLeuArgLys 302
 DB 137 AGTCAACTCGGCTCTGTGTAGCAGCAGGAGGAAAGCGTAAACCGCTCTCAGTCAGTGAAGAA 196
 QY 303 SerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHis 322
 DB 197 ACAGATGGGAACGAGAGAGAGAAATCTGTACAGAGGATA---AGCAGAGCCCATGG 253
 QY 323 GlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThr 342
 DB 254 GGCAATGGCAGCGGAGATCTGTACCCGAGCAAGAGTGTCGTCTGGCTTGGCACA 313
 QY 343 PheGlyThrGlnGluGluAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGly 362
 DB 314 TTCAAAACCGCGCGAAGCTGCTCGAGCCTTACGACGTTGCTGCGATCAAAATCCGTGGC 373
 QY 363 LeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSer 382
 DB 374 CGAAAGCCAACTCAATTTCCCA-----AACACTCAAGTAGAGAGAGAGGCC 421
 QY 383 ProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgProValPro 402
 DB 422 GATATAAACACCGGGGGGAATCAAAATGAGCTGATTTCCGAA-----AAC 466
 QY 403 SerMetMetIleSerAsnValSerGluSerGluAsnSerAlaSerGlyTrpGln 422
 DB 467 CAAGTAGAGACTATCGAGAGACTGATGCGATGGAGGATACATGAGATCTATCAG 506
 QY 423 AsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeuLeuHisGlnGlu 442
 DB 527 ATTCGGTGTGCGCAGCAATCGCGACCGCATATTCGAAATTTATGAGAGTATCAGAC 586
 QY 443 ArgTyrAsnGlyTyrTyrAsnGlyAsnLeuSer 455
 DB 597 TCCAAATTAATCTCTTATTTCCCGCGCGTTTGTCTACT 625

RESULT 11

US-09-964-850-3
 ; Sequence 3, Application US/09964850
 ; Patent No. 6441276
 ; GENERAL INFORMATION:
 ; APPLICANT: Ikeda, Yoshihisa
 ; APPLICANT: Chua, Nam-Hai
 ; TITLE OF INVENTION: ESR2 - A PLANT GENE THAT CAN PROMOTE PLANT REGENERATION AND
 ; TITLE OF INVENTION: TRANSFORMATION
 ; FILE REFERENCE: 2312-111
 ; CURRENT APPLICATION NUMBER: US/09/964,850

; CURRENT FILING DATE: 2001-09-28
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 1265
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (76)..
 ; OTHER INFORMATION: (1059)
 US-09-964-850-3

Alignment Scores:
 Pred. No.: 86-07 Length: 1265
 Score: 154.50 Matches: 80
 Percent Similarity: 38.22% Conservative: 40
 Best Local Similarity: 25.48% Mismatches: 111
 Query Match: 4.95% Indels: 84
 DB: 4 Gaps: 12

US-09-980-364-2 (1-579) x US-09-964-850-3 (1-1265)

QY 251 LysGluGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyThrThr 270
 DB 73 AAAATCGAAAGACCTTGAGAACTTC-----ACC 102
 QY 271 ThrThrThrAsnPhe-----ProMetSerGluTyrGluLysGlu----- 283
 DB 103 GAATCTACCCACTCACCAGACCTAATCTCTCACAATAATTTCTCACCACCTACAGCC 162
 QY 284 -----ValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArg 301
 DB 163 TCACCTGTTAGCGGCAACCGMAACTGCTTCAAAAGATACCACTGTACCATCCTCCCGA 222
 QY 302 LysSerSerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGln 321
 DB 223 GCTGGCAGCAGCAGCAGCAGG-----TACCGCGGCGTA---CGCGGAGGCGC 267
 QY 322 HisGlyArgTrpGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGly 341
 DB 268 TGGGACAGATACCGCGGAGATACGTGACCCCATGTCAAGAGGAGAGACGTGGCTCGGA 327
 QY 342 ThrPheGlyThrGlnGluAlaGluAlaTyrAspIleAlaAlaIleLysPheArg 361
 DB 328 ACATTTGACCGCGGAAACACCGCTTGTGCTTACGACTCTGGCGCTCGTCTTCTCT 387
 QY 362 GlyLeuThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGlu 381
 DB 388 GGAGCAAAAGGCTCGTACTAATTT-----ACTTATCCGACAGCTGTCTATATGCT 438
 QY 382 SerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAlaAsnArgProVal 401
 DB 439 GAACCAAGGTTTCTTTT---TCCAAACAAGAAATCTTCGCCGCTCTGCTGCTCTCT 495
 QY 402 ProSerMetMetIleSerAsn----- 409
 DB 496 CCTTCTCTACCGTTAGATTCTCTTACCCAAACTTTTTCGGTGCACCGCAGCGAGG 555
 QY 410 -----AsnValSerGluSerGluAsnSerAlaSer 419
 DB 556 ATCTATAATACAGCTCTATCTTTTACGCGCGCTGTGTCTCTCTCTCTCTCTCTCT 615
 QY 420 GlyTrpGlnAsnAla-AlaValGlnHisGlnGlyValAspLeuSerLeuLeuHisGln 439
 DB 616 CCTATAATAACTCTTTCAACGGCTCATCAT-----CTTCTTACTCA 657
 QY 439 HisGlnGluArgTyrAsnGlyTyrTyrTyrAsnGlyGlyAsnLeuSer----- 456
 DB 658 GCATCGAATAACCGGATGGGTTTCTTATTCGAAACGAAACACGAGTCTTTTCCC 716
 QY 457 -----GluSerAlaArgAlaCysPheLysGlnGluAspGln----- 469
 DB 717 GGAAGAATCTTCTGATCTGCTATTACAGAGGTCGTTCAAGAGTCTTCTTCAAGAAAAA 776


```

Qy 470 -----HisHisPheLeuSe 474
Db 777 TCGGGGGTTCCTCTCTCCACACACACCGCGGTGACTAGCCATCATGACCACTC 836
Qy 474 rAsnThrGlnSerLeuMetThrAsnIleAspHisGlnSerSerValSerAspSerVa 494
Db 837 TGGTATTCTCTGCTCTCACT---ATATACTCTGAATAATATGTTCAAGAGACTAAGGA 893
Qy 494 lThrValCysGlyAsnValValGlyTyrGlyTyrGln 507
Db 894 GACTTTGCTCGAAACTAGATCGCTACGGGAATTTTCAA 933

RESULT 12
US-09-533-029-11
; Sequence 11, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddle, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 1577
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G4
US-09-533-029-11

Alignment Scores:
Pred. No.: 1.85e-06 Length: 1577
Score: 152.50 Matches: 70
Percent Similarity: 42.75% Conservative: 48
Best Local Similarity: 25.36% Mismatches: 110
Query Match: 4.89% Indels: 49
DB: Gaps: 11

US-09-980-364-2 (1-579) x US-09-533-029-11 (1-1577)
Qy 167 AsnSerSerThrSerCysAspAsnAsnAsnAspSerAsnAsnVal-----182
Db 273 GACTCGGCTTTGACTCGGAAGACGATGATGCTCTTCGTCATGTTAAGCCTTCGTC 332
Qy 183 ---ValAlaGlnGlyThrIleAspAspSerValGluAlaThr-----196
Db 333 TTCACCCAACTACTAAGCCCGTAGCTTCGCTTCCTCCACTGTAGTTCAGCATAT 392
Qy 197 ProLysLysThrIleGluSerPheGlyGln-----ArgThr 208
Db 393 GCCAAGAAACTGTAGAGTCGCTGAGCAAGCTGAGAAATCTTCTAAGAGAGAGGAAG 452
Qy 209 SerIleTyrArgGlyValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTrp 228
Db 453 AATCAATACCGAGGATTAGCGACCGTCTTGG---GGAAATGGGCTCGGAGATCCGT 509

```

```

Qy 229 AspAsnSerCysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGly 248
Db 510 GAT-----CCGAGAAAGGCTCCCGAGAAATGGCTTGGAAACA 545
Qy 249 TyrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGly 268
Db 546 TTCGACATCTGTAGAGACGACAGAGCTTATGATGCTGACGACGCGAGATCCGTGGC 605
Qy 269 ThrThrThrThrThrAsnPheProMetSerGlyTyrGluLysGluValGluGluMetLys 288
Db 606 ACAGAAAGCTAAGTGAATTTCCCGAGGAGAGAACCTAGCTGCTATCCAGAGAAAGCT 665
Qy 289 HisMetThrArgGlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPheSerArg 308
Db 666 CCTAGTGTAAAGACTAAT-----AATCTTCAGAAATCAGTGGCTAAACCAACAAA 716
Qy 309 GlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHisGlyArgTyrGlnAlaArg 328
Db 717 AGCGTAACCTTTGGTTCAGCAGCAACACATCTGAGTCAG-----755
Qy 329 IleGlyArgValAlaGlyAsnLys---AspLeuTyrLeuGlyThrPheGlyThrGlnGlu 347
Db 756 -----CAGTACTGCAACAACCTCTTTGCAACTCTTTTGGTGATATGAGTTTCATGGA 809
Qy 348 GluAlaAlaGluAlaTyrAsp-IleAla---AlaIleLysPheArgGlyLeuThrAlaVa 366
Db 810 GAGAGCCTCAGATGTACAACAATCAGTTGGGTTAACAACCTCGTTCGATGCTGGAGT 869
Qy 366 lThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuPr 386
Db 870 AACATGGATACCACTATTCAGTTCAGTTCAGGCGAGTAACCTCTCTGACTGTCTCTGAG 929
Qy 386 oileGlySerAlaAlaLysArgLeuLysGluAlaAsnArgProValProSerMetMetWe 406
Db 930 TTCGGTGGAGTGATCAGCGGCTTAAAA-----CACCGAG-ATCTCTTCAATGCT 979
Qy 406 rIleSerAsnAsnValSer-----GluSerGluAsnSerAla 418
Db 980 TGTCAATAACAACGAAGCATCATTTGTTGAAGAAACCAATGACGCC 1025

RESULT 13
US-09-533-029-21
; Sequence 21, Application US/09533029
; Patent No. 6664446
; GENERAL INFORMATION:
; APPLICANT: Heard, Jacqueline
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Keddle, James
; APPLICANT: Pineda, Omaira
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Reuber, Lynne
; TITLE OF INVENTION: DISEASE-INDUCED POLYNUCLEOTIDES
; FILE REFERENCE: MBI-010
; CURRENT APPLICATION NUMBER: US/09/533,029
; CURRENT FILING DATE: 2000-03-22
; EARLIER APPLICATION NUMBER: 60/125,814
; EARLIER FILING DATE: 1999-03-23
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1055
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G19
US-09-533-029-21

```

```
Alignment Scores:
Pred. No.: 1.89e-06 Length: 1055
Score: 150.00 Matches: 68
Percent Similarity: 40.61% Conservative: 38
Best Local Similarity: 26.05% Mismatches: 105
Query Match: 4.81% Indels: 50
DB: 4 Gaps: 10

US-09-980-364-2 (1-579) x US-09-533-029-21 (1-1055)
QY 240 LysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGluLysAlaAlaArgAlaTyr 259
DB 121 AAGGCCGTAAACTACCGCTGAGAACTCTGGTCAGAGCTCGATGCTTCGCGCGCGAC 180
QY 260 AspLeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrAsnPhe---ProMetSer 278
DB 181 GAC-----TTCTGGGTTCTATTCCACCTCCAAACTCCATCCCAACCAAC 225
QY 279 GluTyrGluLysGluValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSer 298
DB 226 CAAGTTAACTGAAGAGGAGCGCACTGAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 285
QY 299 LeuArgArgLysSerSerGlyPheSerArgLysAlaSerIleTyrArgGlyValThrArg 318
DB 286 AGGAAGAGGAG-----AATGTTTATAGGGGATA---CGT 318
QY 319 HisHisGlnHisGlyArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeu 338
DB 319 AAGCGTCCATGGGAAATGGCGGCTGAGATTCGAGATCCAGCAAGAAAGGTGTAGAGTT 378
QY 339 TyrLeuGlyThrPheGlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaIle 358
DB 379 TGGCTGTGTAGTTCAACACGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 438
QY 359 LysPheArgGlyLeuThrAlaValThrAsnPhe---AspMetAsnA-gTyrAsnValLys 377
DB 439 CAGATCCGTGTGATAAAGCAAGCTCAACTTCCAGATCTGCACCATCTCCTCCTCCT 498
QY 378 AlaIleLeuGluSerProSerLeuProIleGlySerAlaAlaLysArgLeuLysGluAla 397
DB 499 AATTATATCTCTCCCGCTCATCGCCA-----CGATCAACC 534
QY 398 AsnArgProValProSerMetMetMetIleSerAsnAsnValSerGluSerGluAsnSer 417
DB 535 GATCAGCT---CGGCGAAGAGGTCTCGTGTCTCTCAGAGTGAGAGCGAGTTAAGT 591
QY 418 AlaSerGlyTyrGlnAsnAlaAlaValGlnHisGlnGlyValAlaAspLeuSerLeuLeu 437
DB 592 CAGCCGAGTTTCCCGGTGAGGTATAGGATTTGGAAATGGGACGAGTTTCAG----- 645
QY 438 HisGlnHisGlnGluArgTyrAsnGlyTyrTyrTyrAsnGlyGlyAsnLeuSerSerGlu 457
DB 646 -----AACCTGAGTTACGGA 660
QY 458 SerAlaArgAlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGln 477
DB 661 -----TTTGACCGGATTATGATCTGAAACAGCAGATATCAGCTTCGAA 705
QY 478 SerLeuMetThr-----AsnIleAspHisGlnSerSerValSerAspSerVal 494
DB 706 TCGTTCCTTGAGCTGGAGCGGTAAACAGCGGAGGAGCAACCGAGTTCAGTTGATCGCTT 765
QY 495 Thr 495
DB 766 TCC 768

RESULT 14
US-09-964-850-5
; Sequence 5, Application US/09964850
; Patent No. 6441276
; GENERAL INFORMATION:
; APPLICANT: Ikeda, Yoshihisa
```

```
APPLICANT: Chua, Nam-Hai
; TITLE OF INVENTION: ESR2 - A PLANT GENE THAT CAN PROMOTE PLANT REGENERATION AND
; FILE OF INVENTION: TRANSFORMATION
; FILE REFERENCE: 2312-111
; CURRENT APPLICATION NUMBER: US/09/964,850
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(984)
; US-09-964-850-5

Alignment Scores:
Pred. No.: 1.87e-06 Length: 987
Score: 149.50 Matches: 79
Percent Similarity: 37.17% Conservative: 47
Best Local Similarity: 23.30% Mismatches: 119
Query Match: 4.79% Indels: 95
DB: 4 Gaps: 11

US-09-980-364-2 (1-579) x US-09-964-850-5 (1-987)
QY 253 GluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyThrThrThr 272
DB 4 GAAAGAAGCTTGAGAACTTC-----ACGATCT 33
QY 273 ThrAsnPhe-----ProMetSerGluTyrGluLysGlu----- 283
DB 34 ACCACTCACCAGACCTTAATCTCTCACAATAATCTTCACTGAACCTACAGCCTCACCT 93
QY 284 ValGluGluMetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgLysSer 303
DB 94 GTTAGCGCAACCGCAAACTGTCTTCAAAAGATACCATGTGAACCATCGCGGAGCTGGC 153
QY 304 SerGlyPheSerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHisGly 323
DB 154 AGCAGCACACGAGG-----TACCGCGCGTA---CGCGGAGGCGCGTGGGA 198
QY 324 ArgTyrGlnAlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPhe 343
DB 199 CGATACGCGCGAGATACGTGACCAATGTGAAAGAGAGACGCTGGCTCGGAACATTT 258
QY 344 GlyThrGlnGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeu 363
DB 259 GACACGCGCAACAGCGCTGTGCTTACGACTCTGGGCTCGTCCCTTCGTGGAGCA 318
QY 364 ThrAlaValThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSer--- 382
DB 319 AAGGCTCGTACTAATTT-----ACTTATCGCAGAGCTGTCAATTATGCTGAACCA 369
QY 383 -----ProSerLeu 385
DB 370 AGTTTCTTTTCCAAAGAAATCTTCGCGCTGTCTGCTGCTCTCTCTCTCTCTCTCT 429
QY 386 ProIleGlySer-----AlaAlaLysArg----- 393
DB 430 CGTTAGATTCTCTTACCACAAATTTTACGGTGACCGGACGAGCAGAGGATCTATAAT 489
QY 394 -----LeuLysGluAlaAsnArgProValProSerMetMetMetIleSer 408
DB 490 ACACAGTCTATCTTACGCGACGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 549
QY 409 AsnAsnValSerGluSerGluAsnSerAlaSerGlyTyrGlnAsnAlaAlaValGlnHis 428
DB 550 AACTCTTTCAACGGCTCATCTCTTACTCAGCATCGAAACCGCATGGTTCCTTCTTAT 609
QY 429 HisGlnGlyValAsp-LeuSer----- 435
```

```
Db 610 TCCGAAACGAAACAAACACGAGTCGTTTTTCCCGGAAGAAATCTTCTGATACTCGTCTATTA 669
QY 436 -----LeuLeuHisGlnHi 440
Db 670 CAAGAGTCGTTCAAGAGTCCTTGAAGAAAAATCCGGCGTTCCTCCCTCTCCACCAACA 729
QY 440 sGlnGluArgTyrAsnGlyTyrTyrAsnGlyGly-----AsnLeuSerSerG1 457
Db 730 CCACCCCGCGTG-ACTAGCCATCATGACAACTCTGGTTATTTCTCTAATCTCCTCATATATA 788
QY 457 uSerAlaArgAlaCysPheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrG1 477
Db 789 CTTCTGAATATCGTTCAAGAGACTTAAGGAGACTTTGTCTCGAAACTAGATCGCTACGG 848
QY 477 nSerLeuMetThrAsnIleAspHisGlnSerSerValSerAspSerValThrValCy 497
Db 849 GAATTTTCAAGCTAATGACGACCGCGTAAGACCGCTCGCAGACGGTGTATTATCTGTTGGG 908
QY 497 sGlyAsnValValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCys 515
Db 909 ATCAACGAGTGGGGGTATCAAGAAATGTTGATGACGGAATCAGTTAGGCTGT 963
```

RESULT 15

```
US-09-202-161B-6
; Sequence 6, Application US/09202161B
; Patent No. 6653533
; GENERAL INFORMATION:
; APPLICANT: Purdue Research Foundation
; TITLE OF INVENTION: GENES ENHANCING DISEASE RESISTANCE IN PLANTS
; FILE REFERENCE: 7024-371
; CURRENT APPLICATION NUMBER: US/09/202,161B
; CURRENT FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: PCT/US97/10382
; PRIOR FILING DATE: 1997-06-12
; PRIOR APPLICATION NUMBER: 60/046,494
; PRIOR FILING DATE: 1997-05-14
; PRIOR APPLICATION NUMBER: 60/019,633
; PRIOR FILING DATE: 1996-06-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: ASCII
; SEQ ID NO 6
; LENGTH: 1405
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-09-202-161B-6
```

Alignment Scores:

```
Pred. No.: 9.83e-06 Length: 1405
Score: 145.00 Matches: 50
Percent Similarity: 46.55% Conservative: 31
Best Local Similarity: 28.74% Mismatches: 67
Query Match: 4.65% Indels: 26
DB: 4 Gaps: 7
```

US-09-980-364-2 (1-579) x US-09-202-161B-6 (1-1405)

```
QY 144 AsnGlnProValAspAsnValAspAsnGlnGlu---AsnGlyAsnAlaAlaLysGlyLeu 162
Db 366 AACAAAGCATGTTTCTGAGCATACGAAATAATCCAAATTCAGATTACAAAGAGTTGTG 425
QY 163 SerLeuSerMetAsnSerThrSerCysAspAsnAsnAspSerAsnAsnVal 182
Db 426 AGGATTATCTTACAGATCCGATGCTACAGATTCTTCGGATTCAGAGCCGGAATACT 485
QY 183 ValAlaGlnGlyLysThrIleAspAspSerValGluAlaThrProLysLysThrIleGlu 202
Db 486 GTACGGAGATGAAGAGCCAGCTGACGGAGATCAACCTTATGCG-----TCAACCAAA 539
QY 203 SerPheGly-----GlnArgThrSerIle----- 210
Db 540 TCGATCGCGCACGAAACACGAGATCGGTCTCCGGATTCTCGAGTCTCAGCTCGTGGAAA 599
QY 211 ---TyrArgGlyValThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTyrAsp 229
```

```
Db 600 AAGTTTATAGAGCGGTTCTCAAGAACCCGTGG---GGTCGTTGGGCTGCAGAGATTCCGGAC 656
QY 230 AsnSerCysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyr 249
Db 657 CCGACC-----CGGGAAACCGGTGTGTTGGGTACTTAT 692
QY 250 AspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyThr 269
Db 693 GACACCCCAAGAAAGACGCTGCTGTTTACGATAAAGCTCAGTTAAAGCTCAAAGGTCCT 752
QY 270 ThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGluMetLysHis 289
Db 753 GACGCCGTTACCAATTTTCCGGTATCAACA---ACGCCGGAGGTAAACGGTACCGTTTACG 809
QY 290 MetThrArgGlnGluTyrValAlaSerLeuArgArgLysSer 303
Db 810 GAAACCCGAAACCGAGTCTGTTGCCGACGGTGGAGATAAAAGC 851
```

Search completed: July 19, 2004, 18:30:25
Job time : 164 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 19, 2004, 17:20:03 ; Search time 742 Seconds
(without alignments)
3806.904 Million cell updates/sec

Title: US-09-980-364-2
Perfect score: 3119
Sequence: 1 MNNNLGFLSPYEQNHRRK.....YYHGEGGCVAPFTTWNND 579

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3190992 seqs, 2439311697 residues

Total number of hits satisfying chosen parameters: 6381984

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0980364/runat_15072004_095239_25984/app.query.fasta_1.775
-DB=published Applications NA -QFMT=fastap -SUPER=rnpb-MINMATCH=0.1
-LOFPCLO=0 -LOFPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdd -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0980364@cgn 1 1 516 @runat_15072004_095239_25984
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSEBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
------------	-------------	-------	--------	----	-------------

ALIGNMENTS

RESULT 1

US-10-183-687-351
; Sequence 351, Application US/10183697
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B01458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27

1	1310	42.0	2588	13	US-10-183-687-351	Sequence 351, App
2	1296	41.6	2463	13	US-10-183-687-355	Sequence 355, App
3	1119.5	35.9	1990	13	US-10-183-687-361	Sequence 361, App
4	1111.5	35.6	2169	13	US-10-424-599-91337	Sequence 91337, A
5	1110.5	35.6	2088	16	US-10-374-780A-1738	Sequence 1738, App
6	1075.5	34.5	1824	13	US-10-183-687-353	Sequence 353, App
7	1047.5	33.6	1818	13	US-10-225-066A-365	Sequence 365, App
8	1047.5	33.6	1818	13	US-10-225-066A-855	Sequence 855, App
9	1047.5	33.6	1818	13	US-10-225-066A-911	Sequence 911, App
10	1047.5	33.6	1818	16	US-10-374-780A-333	Sequence 333, App
11	1017	32.6	1935	13	US-10-183-687-486	Sequence 486, App
12	1003	32.2	1959	13	US-10-183-687-484	Sequence 484, App
13	1000	32.1	1878	13	US-10-183-687-482	Sequence 482, App
14	996.5	31.9	1954	13	US-10-183-687-490	Sequence 490, App
15	996	31.9	2116	13	US-10-424-599-100930	Sequence 100930,
16	987.5	31.7	1874	13	US-10-183-687-347	Sequence 347, App
17	986.5	31.6	2128	16	US-10-374-780A-1763	Sequence 1763, Ap
18	979	31.4	2374	13	US-10-183-687-488	Sequence 488, App
19	977	31.3	2052	13	US-10-225-066A-1013	Sequence 1013, Ap
20	977	31.3	2052	16	US-10-374-780A-2495	Sequence 2495, Ap
21	974.5	31.2	1941	13	US-10-225-066A-381	Sequence 381, App
22	974.5	31.2	1941	16	US-10-374-780A-345	Sequence 345, App
23	973.5	31.2	1926	14	US-10-024-632-545	Sequence 545, Appli
24	971.5	31.1	1920	17	US-10-437-963-49245	Sequence 49245, A
25	971	31.1	2168	13	US-10-183-687-337	Sequence 337, App
26	964.5	30.9	1758	13	US-10-183-687-337	Sequence 337, App
27	960	30.8	2147	13	US-10-424-599-70525	Sequence 70525, A
28	957.5	30.7	1290	13	US-10-183-687-359	Sequence 359, App
29	956.5	30.7	1985	13	US-10-435-114-783	Sequence 783, App
30	949.5	30.4	1753	13	US-10-183-687-333	Sequence 333, App
31	948	30.4	794	13	US-10-425-114-20762	Sequence 20762, A
32	948	30.4	806	16	US-10-374-780A-1742	Sequence 1742, Ap
33	948	30.4	1479	13	US-10-425-114-14859	Sequence 14859, A
34	948	30.4	1864	13	US-10-183-687-331	Sequence 331, App
35	947.5	30.4	1466	13	US-10-425-114-3280	Sequence 3280, Ap
36	944.5	30.3	2102	17	US-10-437-963-37126	Sequence 37126, A
37	942	30.2	2010	14	US-10-024-632-8	Sequence 8, Appli
38	942	30.2	2344	14	US-10-024-632-1	Sequence 1, Appli
39	941.5	30.2	1830	17	US-10-437-963-20346	Sequence 20346, A
40	941	30.2	1500	13	US-10-225-066A-591	Sequence 591, App
41	941	30.2	1500	16	US-10-374-780A-399	Sequence 399, App
42	935.5	30.0	1182	13	US-10-425-114-19324	Sequence 19324, A
43	924	29.6	2323	14	US-10-024-632-3	Sequence 3, Appli
44	913.5	29.3	1738	15	US-10-059-911-4	Sequence 4, Appli
45	901.5	28.9	1869	17	US-10-437-963-84937	Sequence 84937, A

[illegible]

RESULT. 3

RESULT 3
US-10-183-687-361
; Sequence 361, Application US/10183687
; Publication No. US20030204870A1

; FUDICATION NO: 0320
; GENERAL INFORMATION:

GENERAL INFORMATION:

APPLICANT: Allen, Steve
APPLICANT: Allen, William B.
APPLICANT: Cahoon, Rebecca
APPLICANT: Epelbaum, Sabine
APPLICANT: Famodu, Omolayo O
APPLICANT: Harvell, Leslie T
APPLICANT: Jones, Todd
APPLICANT: Kinney, Tony
APPLICANT: Kleim, Ted

215 ThrArgHisArgTrpThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysIysArg 234
Db ACTCGACATAGATGGAGTGGAGGTATGAAGCTCATCTTTGGGATAATAGCTGTAGAAGG 689
235 GluGlyGlnThrArgIysGlyArgGlnValTyrLeuGlyGlyTyrAspIysGlnGluIys 254
Db GAAGGCCAATCAAGAAAGAGCCCAAGTTTATTTGGGTGGATATGATTAAGAGAGAAA 749
255 AlaAlaArgAlaTyrAspLeuAlaAlaLeuIysTyrTrpGlyThrThrThrThrAsn 274
Db GCAGCTAGAGCTTATGATTAGCTGACCTGAAGTACTGGGGGACATCCACCCTACCAAC 809
275 PheProMetSerGluTyrGlnIysGluValGluGluMetIysHisMetThrArgGlnGlu 294
Db TTTCCAAATAGCAACTATGAGAGGAATTTGGATGAATCAACACATGACGAGCAAGAA 869
295 TyrValAlaSerLeuArgIysSerSerGlyPheSerArgGlyValaSerIleTyrArg 314
Db TTTGTTGCCGCCATTAGAGAAAAGCAGTGGTTCTCCAGGGGTCATCAATGATCGT 929
315 GlyValThrArgHisGlnHisGlyArgTrpGlnAlaArgIleGlyArgValAlaGly 334
Db GGAGTTTACAAGGCATCACCAACACGGAAGATGCAAGCAAGGATTGGCAGAGTTGCAGA 989
335 AsnIysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAsp 354
Db AACAAAGATCTTTACTTTGGAACTTTTCAGTACTGAGGAAGAGCTGCAGAGCATACGAC 1049
355 IleAlaAlaIleIysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyr 374
Db ATAGCAGCGATAAAGTTCAAGGCTCAACGCTGTCAACAACTTTGACATGAGCCGCTAC 1109
375 AsnValIysAlaIleLeuGluSerProSerLeuProIle---GlySerAlaAlaIysArg 393
Db GACGTGAAGCCATTCTTGAAGCAACACTCTCCCAATGAGGAGGCGCTGCAAGCGT 1169
394 LeuIysGluAla-----AsnArgProValProSerMetMetIle--- 407
Db CTGAAGAAGCTCAAGCTCAGAACTCTTCGAGAAAACGCGAAGAGATGATGCTACTAGGC 1229
408 SerAsnAsnValSerGluSerGluAsnSerAlaSerGlyTrpGlnAsnAlaAla----- 425
Db TCATCTTCCAGTTCCCAATACGGAACCTCAGCAAGCTCTTCTAGGCTTCACGCTTACCCT 1289
426 ---ValGlnHisHis-----GlnGlyVal 432
Db CTAATGAGCAGCACCAACAGTTCGAGCAACCTCACTCTGCTAACTCTTCAAAACCAC 1349
433 AspLeuSerLeuLeuHis---GlnHisGlnGlu-----ArgTyrAsnGlyTyrTyr 448
Db GACATAAGTTCTTCTCACTTCTCTCACCAGCAAGACCTTTGTCATCATCAGGGTTACATC 1409
449 TyrAsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCysPheIys---GlnGluAsp 467
Db CAAACGAGCTTCAGTTGCACAGCAGAGTGGCGCTTCTTCTATAGCTTTTCAAGATAAT 1469
468 AspGlnHisHis-----PheLeuSerAsnThrGlnSerLeu-----MetThrAsn 482
Db GCTCAGTTCTACAATAGTTTACCTTCAGAACCCCTGCAATGCTTCAGGGAATGATGAAC 1529
483 IleAspHisGlnSerSerValSer-----AspAspSerValThrValCys 497
Db ATGGGGTCTTCTTCTTCTTCTCTGTTGAGAAATATATATAGTAACATAATAT 1589
498 GlyAsnValValGlyTyrGlyGlyTyrGlnGlyPheAlaAlaProValAsnCysAspAla 517
Db AATAATGTTGGTGGTTTGTGGGA---AGTGGGTTTGGTATGGCTTCCAAATGCAACGGCG 1646
518 -----TyrAlaAlaSerGluPheAspTyrAsn 526
Db GGGAAACCGTGGGACAGCGAGGAGTTAGGGCTGGTGGAGGTGGAGTATGACATGCCG 1706
527 AlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGlnGlnSerProGlyGlyAsp 546

Db 1707 GCTGGAGGTTACGGTGGCTGGTGGCGCGGACTCCATGCAGACGTCAAATGGTGGG--- 1763
Qy 547 PheProAlaAlaMetThrAsnAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGly 566
Db 1763 ----- 1763
Qy 567 GlyGluValAlaProThrPheThrValTrpAsnAsp 578
Db 1764 -----GTGTTCACAATGTGAATGAT 1784
RESULT 4
US-10-424-599-91337
; Sequence 91337, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Wihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; CURRENT FILING DATE: 2003-04-28
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 91337
; LENGTH: 2169
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53488C.1
US-10-424-599-91337

Alignment Scores:
Pred. No.: 2,628-119 Length: 2169
Score: 1111.50 Matches: 274
Percent Similarity: 54.01% Conservatives: 76
Best Local Similarity: 42.28% Mismatches: 142
Query Match: 35.64% Indels: 157
DB: 13 Gaps: 20

US-09-980-364-2 (1-579) x US-10-424-599-91337 (1-2169)
Qy 1 MetAsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgIys 20
Db 133 ATGAACAACAACCTGGCTTTCGTTCCCTCTTCT----- 165
Qy 21 AspValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
Db 165 ----- 165
Qy 41 ProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSerPheProSerProPheGly 60
Db CCTACTCATTTCTTCTTACCAGCTCATGATCTTCAAGCAACTCAATATCATCAATTTTCC 225
Qy 61 ValValValAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTyrAsp---Ile 79
Db CTTGGGTTAGTGAACGAGAACATGATGATACCCCTTTCCAAATCATGATTTGGATCTGAT 285
Qy 80 AsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProIysLeuGluAsn 99
Db AACACCCATAGTACCAACGAAAT-----CCAAAGTGGCTGAT 324
Qy 100 PheLeuGlyArgThrThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGly 119
Db TTTCTAGGA-----GTGAGCAAGTCTGAAATCAGTCAGACCTTCGACCC 369
Qy 120 SerGlyCysTyrGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMetIleIys 139
Db TTAACGAATTCATTCAATGATTCAGATTATCTGTTTCAACAAACAGCTGTGGCT 429
Qy 140 ThrTrpLeuArgAsnGlnProValAspAsnValAspAsn-----GlnGluAsnGlyAsn 157


```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1738
; LENGTH: 2088
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Predicted polypeptide sequence is orthologous to G1793
US-10-374-780A-1738

Alignment Scores:
Pred. No.: 3,23e-119 Length: 2088
Score: 1110.50 Matches: 274
Percent Similarity: 53.86% Conservative: 75
Best Local Similarity: 42.28% Mismatches: 139
Query Match: 35.60% Indels: 161
DB: 16 Gaps: 20

US-09-980-364-2 (1-579) x US-10-374-780A-1738 (1-2088)

QY 1 MetAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLys 20
Db 133 ATGAACACCACTGGCTTCGTTCCCTCTTCT- 165
QY 21 AspValTyrSerSerThrThrThrThrValValAspValAlaGlyGluTyrCysTyrAsp 40
Db 165 ----- 165
QY 41 ProThrAlaAlaSerAspGluSerAlaIleGlnThrSerPheProSerProPheGly 60
Db 166 CCTACTCATCTTCCTTACCAGCTCATGCTTCACGCACTCAATCATCAATTTCC 225
QY 61 ValValValAspAlaPheThrArgAspAsnSerHisSerArgAspTrpAsp 79
Db 226 CTTGGGTAGTGAACGAGAACATGATACCCCTTCCAAATCATGATTGGAATCTGATT 285
QY 80 AsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAsn 99
Db 286 AACCCCATAGTACGACGAAAT- 324
QY 100 PheLeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGly 119
Db 325 TTCTTAGGA-----GTGACCAAGTCTGAAATCAGTACAGACCTTGCAGCG 369
QY 120 SerGlyCysTyrGlyGlyAspGlyGlyGlySerLeuGlyLeuSerMetIleLys 139
Db 370 TTAACGAAATTCATCAATGATTCAGATATCTGTTCACAAACACAGCTCTGGTGCCT 429
QY 140 ThrTrpLeuArgAsnGlnProValAspAsnValAspAsn-----GlnGluAsnGlyAsn 157
Db 430 -----ATGCAAAACCCCTGCTTGGACACACCTAGCAATCAGTATCAAGAAATGCTAAT 483
QY 158 Ala---AlaLysGlyLeuSerLeuSerMetAsnSer-----SerThrSerCysAspAsn 174
Db 484 AGTAATTTGCAATCATTCACATATCCATGGGAGTGGTAAGGATTCAACATGTGAAC 543
QY 175 AsnAsnAspSerAsnAsnValValAlaGlnGlyLysThrIleAspAspSerValGlu 194
Db 544 AGTGGTGAATATACCAACACT- 576
QY 195 AlaThrProLysLysThrIleGluSerPheGlyGlnArgThrSerIleTyrArgGlyVal 214
Db 577 GTTGACCTCAAGAACTTTGGATACATTCGGGAGAGAACATCCATATATCTGTGGAGTA 636
QY 215 ThrArgHisArgTyrThrGlyArgTyrGluAlaHisLeuTrpAspAsnSerCysLysArg 234
Db 637 ACTGCACATAGTACTGGAGGTATGAAGCTCATCTTTGGGATATATAGCTGTAGAGG 696
QY 235 GluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLysGluGluLys 254
Db 697 GAAGCCCAATCAAGAAAGGACGCCAAGCTTATTGCTGGTATATGATATAAGAGAGAAA 756
QY 255 AlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTrpGlyThrThrThrThrAsn 274

```

```

757 GCAGCTAGGGCTTATGATTTAGCTGACTGAAGTACTGGGGACATCCACACACTACCAAC 816
QY 275 PheProMetSerGluTyrGluLysGluValGluMetLysHisMetThrArgGlnGlu 294
Db 817 TTTCCAAATTAGTAATAAGAGGATTCGATGAATGAACACATGACGCGACAGAA 876
QY 295 TyrValAlaSerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSerIleTyrArg 314
Db 877 TTTGTTGCTGCCATTAGAGGAAAGCAGTGGTTTCTCCAGGGGTGCATCAATGTATCGT 936
QY 315 GlyValThrArgHisHisGlnHisGlyArgTyrGlnAlaAlaArgIleGlyArgValAlaGly 334
Db 937 GGAGTTACAGGCAATCACCAACACGGAAGATGGACACAGAAATTCGACAGTTGACAGGA 996
QY 335 AsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluGluAlaAlaGluAlaTyrAsp 354
Db 997 AACAAAGATCTTTACTTTGGGAACCTTCAGTACTGAAGAAGAGGCTGCTGAAGCATACGAC 1056
QY 355 IleAlaAlaIleLysPheArgGlyLeuThrAlaValThrAsnPheAspMetAsnArgTyr 374
Db 1057 ATAGCTGCGATAAAGTTTCAGAGGCTCAACGCTGTCACAAACTTTTGACATGAGCGGTAC 1116
QY 375 AsnValLysAlaIleLeuGluSerProSerLeuProIle---GlySerAlaAlaLysArg 393
Db 1117 GACGTGAAGGCCATCTTGAAGCAACACTCTCCCAATAGGAGGAGCTGCAAGCGT 1176
QY 394 LeuLysGluAla-----AsnArgProValProSerMetMetMetIleSer 408
Db 1177 CTGAAGAAGCTCAAGCTCTAGAATCTTCGAGAAGCGCGAAGAGATGATTCACATAGGA 1236
QY 409 AsnAsnValSerGluSerGluAsnSerAlaSerGlyTrpGlnAsnAlaAla----- 425
Db 1237 TCATCCACATTCCTCAATATGGAACCAACAGCTCTAATTCAGGCTACATGCTTACCTCTA 1296
QY 426 ValGlnHisHis-----GlnGlyValAsp 433
Db 1297 ATGACGACCAACCAACCACTGCTGCTCAACCTGCTTAACCTTTCGAAACCATGAT 1356
QY 434 LeuSerLeu-----LeuHisGln----- 439
Db 1357 ATCAGTTCTCCTCTCTCCACGACAGACCTTTGGCATCAGGTTACATCCAAACGACAG 1416
QY 440 -----HisGlnGluArg----- 443
Db 1417 CTTCACTGGTGCACGACGACAGAGTGGTGGTCTTCTCTTATAGCTTTTCAGAAATAATAT 1476
QY 444 ---TyrAsnGlyTyrTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCys 462
Db 1477 ATAAATAATGCTCAGTCTTATAATGTTATATCTCAGAACCCCTGCTGCTT--- 1533
QY 463 PheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeuMetThrAsn 482
Db 1534 -----CAGGNAATGATTAACTGGGTCTTCATCTTCTTCATCTGTG 1575
QY 483 IleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsnValValGly 502
Db 1576 TTGGAGATAATAATAGTACCAATAATAATGTTGG-TGGGTTTGGGAAGTGGGTTTGG 1634
QY 503 TyrGlyGlyTyrGlnGlyPheAlaProValAsnCysAspAlaTyrAlaAlaSerGlu 522
Db 1635 TATGGC----- 1640
QY 523 PheAspTyrAsnAlaArgAsnHisTyrTyrPheAlaGlnGlnGlnThrGlnGlnSer 542
Db 1641 TTCTAATGCAACGCTCGGGGAACAC----- 1664
QY 543 ProGlyGlyAspPheProAlaAlaMetThrAsn-AsnValGlySerAsnMetTyrTyrHis 562
Db 1665 ---GGTGGGAGCGCGGAGGAGCTAGGGCTGGTGAAGTGGACTATGACATGCCACTGG 1721
QY 562 sGlyGluGlyGlyGlyGluValAlaPro----- 571
Db 1722 TGGTTACGGTGGATGCTCGCGCGCGCGGAGTCCATGCAGACGTCGATAGTGG 1781

```

QY 572 -ThrPheThrValTrpAsnAsp 578
 DB 1782 GGTGTTTCACATGTGGATGAC 1803

RESULT 6
 US-10-183-687-353
 ; Sequence 353, Application US/10183687
 ; Publication No. US20030204870A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Allen, William B.
 ; APPLICANT: Cahoon, Rebecca
 ; APPLICANT: Epeibaum, Sabine
 ; APPLICANT: Famodu, Omolayo O.
 ; APPLICANT: Harvelli, Leslie T.
 ; APPLICANT: Jones, Todd
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Klein, Ted
 ; APPLICANT: Li, Changjiang
 ; APPLICANT: Oliveira, Igor Cunha
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Shen, Bo
 ; APPLICANT: Tarczynski, Mitchell C.
 ; TITLE OF INVENTION: Alteration of Oil Traits in Plants
 ; FILE REFERENCE: BBI458 US NA
 ; CURRENT APPLICATION NUMBER: US/10/183,687
 ; PRIOR FILING DATE: 2002-06-27
 ; PRIOR APPLICATION NUMBER: 60/301,913
 ; FILING DATE: 2001-06-29
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 353
 ; LENGTH: 1824
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-183-687-353

Alignment Scores:
 Pred. No.: 3,38-115 Length: 1824
 Score: 1075.50 Matches: 240
 Percent Similarity: 59.62% Conservative: 45
 Best Local Similarity: 50.21% Mismatches: 98
 Query Match: 34.48% Indels: 95
 DB: 13 Gaps: 12

US-09-980-364-2 (1-579) x US-10-183-687-353 (1-1824)

QY 167 AsnSerSerThrSerCysAspAsnAsnAspSerAsnAsnValValAlaGlnGly 186
 DB 12 AATGAGAGAGTCTCTCTGATAACAAACCA-----AACACCTCGGCTGCATTGAT 65
 QY 187 LysThrIleAspSerValGluAlaThrProLysLysThrIleGluSerPheGlyGln 206
 DB 66 TCCACCAACCGGAGCATGAACTGCACCCAGAAAGTCCATTGACACTTTTGACAG 125
 QY 207 ArgThrSerIleThrArgGlyValThrArgHisArgTrpThrGlyArgTrpGluAlaHis 226
 DB 126 AGAAGCTTCTATCTACCGTGGTGTAAACAGGCATAGGTGGACGGGAGGTACGAGCTCAC 185
 QY 227 LeuTrpAspAsnSerCysLysArgGluGlnThrArgLysGlyArgGlnValTrpLeu 246
 DB 186 CTGGGGATATAGTTGTAGAGAGAGGAGCAGACTCGCAAGGAAAGGCAAGTTTACTTG 245
 QY 247 GlyGlyTrpAspLysGluGluLysAlaAlaArgAlaTrpAspLeuAlaLeuLysTyr 266
 DB 246 GGTGGTTATGATAAGAGAAAGGACGCTAGAGCTACGATTGGCAGCACTAAATAC 305
 QY 267 TrpGlyThrThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGlu 286
 DB 306 TGGGGAACCAACCAACCAAAATTTTCCAAATTAGCCACTATGAGAAAGAGTTTGAAGAA 365
 QY 287 MetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPhe 306

DB 366 ATGAAGCACATGCTAGGCAAGAGTACGTTGGTCACTTGGAGAGAGAGTAGTGGT 425
 QY 307 SerArgGlyAlaSerIleThrArgGlyValThrArgHisHisGlnHisGlyArgTrpGln 326
 DB 426 TCTCGCGGTGCATCCATTTATCGAGGAGTGAGGAGACACCAACAGAGTGGCAA 485
 QY 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGln 346
 DB 486 CGAGAGATTGGAGAGTCTGCGCAACCAAGATCTTTACTTGGAACTTTTAGCAACCA 545
 QY 347 GluGluAlaAlaGluAlaTrpAspIleAlaAlaIleLysPheArgGlyLeuThrAlaVal 366
 DB 546 GAAGAGCGAGCGAAGCATATGATGAGCAGCAATCAAAATTCGAGGAGTAAAGTCTGT 605
 QY 367 ThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluLysProSerLeuPro 386
 DB 606 ACAAACTTTGACATGAGCAGATATGACGTGAAAGCATACTTGAGAGCACCACCTTTGCCA 665
 QY 387 IleGlySerAlaAlaLysArgLeuLysGlu----- 396
 DB 666 ATAGTGGTGTGCAAGCGTTTGAAGGATATGAGCAGGTTGAACTCAGTGTGGATAAT 725
 QY 397 -----AlaAsnArgProValProSerMetMetMetIleSerAsnAsnValSerGlu 413
 DB 726 GGTATAGAGCAGATCAAGTAGATCATATGATCATGAGTTCTCAGCTAACCAAGGA 785
 QY 414 SerGluAsnSerAlaSerGlyTrpGlnAsnAlaAlaValGlnHisGlnGlyValAsp 433
 DB 786 ATCAATAACAACCTATGCGAGGGGGGAACAGCACT-----CATCATACTGGCACAAT 839
 QY 434 LeuSerLeuLeuHisGlnHisGlnGluArgTyrAsnGlyTyrTyrAsnGlyGlyAsn 453
 DB 840 GCTCATGCAATTCACCAACCTCACTTGCACCACTGCACCTACCTTATGGACAAAGA 899
 QY 454 LeuSerSerGluSerAlaAlaGlyCysPheLysGlnGluAspAspGln----- 469
 DB 900 ATTAAT-----TGGTGAAGCAAGCAACCAAGCAACTCTGATGCCCT 944
 QY 470 -----HisHis 471
 DB 945 CACTCTTTGTCTTATTACAGATATTCACTCACTTCAGCTAGGGAACAATGGAACATAAC 1004
 QY 472 PheLeuSerAsnThrGlnSerLeu-----MetThrAsnIleAspHisGlnSerVal 489
 DB 1005 TTCTTTCACAAAATTCAGGGTTGCACCTATGTTGACATGGAT-----TCTGCT 1055
 QY 490 SerAspAspSerValThrValCysGlyAsnValVal-----GlyTyrGlyTyrGln 507
 DB 1056 TCCATTGACAATAGTCTCTTCTTAACCTGGTGTGTTATGATGTTATGAGGTGGTGG 1115
 QY 508 GlyPhe---AlaAlaProVal---AsnCysAspAlaTrpAlaAlaSerGluPheAspTyr 525
 DB 1116 GCGTCAATATGATGCTTATGAGAACTACTACTGCTGTGTTGCAAGTATGATGGTCAAA 1175
 QY 526 Asn----- 526
 DB 1176 AATCCAAGCAACATCATGGTTTGGTGATAATGAGATAAAGCACTGGTTATGAAAGT 1235
 QY 527 -----AlaArgAsnHisTyrTyrPheAla 534
 DB 1236 GTGTATGGCTCTGCAACTGATCTTATCATGCATGCAAGAACTTGTATTATCTTACT 1285
 QY 535 GlnGlnGlnGlnThrGlnGlnSerProGlyGlyAspPheProAlaAlaMetThrAsnAsn 554
 DB 1296 CAACAGCAA-----TCATCTTCTGTGTATACA 1322
 QY 555 ValGlySerAsnMetTyrTyrHisGlyGluGlyGlyGlyGluValAlaProThr 572
 DB 1323 GTGAAGGCTAGTGCATATGATCAAGGGTCTGCATGCACTACTTGGTTCACAACT 1376

RESULT 7
 US-10-225-066A-365

```

; Sequence 365, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RIECHMANN, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROWN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MB10036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 365
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-10-225-066A-365

Alignment Scores:
Pred. No.: 6.26e-112 Length: 1818
Score: 1047.50 Matches: 262
Percent Similarity: 46.96% Conservative: 70
Best Local Similarity: 37.06% Mismatches: 110
Query Match: 33.58% Indels: 265
DB: 13 Gaps: 23

US-09-980-364-2 (1-579) x US-10-225-066A-365 (1-1818)
QY 2 AsnAsnAsnProLeuGlyPhe-----SerLeuSerProTyrGlu 14
DB 65 TCTAACCACTGGCTGGCTTCCTCTTCCCGAACACTCTCTTGGCTCTCATGAA 124
QY 15 GlnAsnHisHisArgLysAspValTyrSerSerThrThrThrValValAspValAla 34
DB 125 TACAAC-----130
QY 35 GlyGluTyrCysTyrAspProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSer 54
DB 130-----130
QY 55 PheProSerProPheGlyValValAspAlaPheThrArgAspAsnAsnSerHisSer 74
DB 131-----CTTGGCTTGGTCAGCGACCATATG-----GACAAACCTTTTCAAAACA 172
QY 75 ArgAspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGly 94
DB 173 CAAGAGTGAATATG-----ATCAATCCACACCGGTGGAGGAGGATGAA 217
QY 95-----ProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThr 110
DB 218 GGAGAGAGGTTCNAAAGTGGCGATTTCTCGGTGTGACCAACCGGACGAAACCAA 277
QY 111 AsnGluAsnValGlyAspGlySerGlySerGlyCysTyrGlyGlyGlyAspGlyGlyGly 130

```

```

DB 278 TCCAAACCACTAGTAGCTTACAACGACTCAGACTACTACTACTTCCATACCAAT----- 328
QY 131 GlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProVal----- 147
DB 329 -----AGCTTGATCGCTAGCGTCCCAATCAACAGATGCTGTTGTAGCAGCT 373
QY 148 -----AspAsnValAspAsn-----GlnGluAsnGlyAsnAla 158
DB 374 TGTGACTCCATATCTCTTAACAACAGTAGCTATCATGAGCTTCAAGAGAGTGTCTCACAAT 433
QY 159 AlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSer 178
DB 434 CTACAGTCACTACTTTGTCCATGGGACCACC-----GCT 469
QY 179 AsnAsnAsnValVal-----AlaGlnGlyLysThrIleAspAspSerValGlu--- 194
DB 470 GGTATATATGTTGTAGACAAAGCTTCACATCCGAGACCACCGGGGATAACGCTAGCGGT 529
QY 195 -----AlaThrProLysLysThrIleGluSerPheGlyGln 206
DB 530 GGAGCACTAGCGCTTGTAGACGCGCCACGAGCGCTGCAATGGACACTTTCCGACAA 589
QY 207 ArgThrSerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHis 226
DB 590 CGAACCTCGATCTATCGTGGTGTACAGACATCGATGGACTGTCGATATAGGCTCAT 649
QY 227 LeuTrpAspAsnSerCysLysArgGlyGlnThrArgLysGlyArgGlnValTyrLeu 246
DB 650 CTATGGGATAATAGTTGTAGAAAGGAGCGCCAGTCTAGGAAAGGAAGACAAGTTTACTTG 709
QY 247 GlyGlyTyrAspLysGluGlyLysAlaAlaAlaGlyValThrAspLeuAlaAlaLeuLysTyr 266
DB 710 GGTGATATGACAAAGAAGATAAGCAGCAAGATCATATGATCTAGCTGCATTAAGTAC 769
QY 267 TrpGlyThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGlu 286
DB 770 TGGGTCTCTCACTACTACTTAATTTTCCCTACAAACTACGAGAAAGAAAGTAGAGAA 829
QY 287 MetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgLysSerSerGlyPhe 306
DB 830 ATGAAGACATGACAGAGCAAGAGTTCGTGGTGCCTTAGAAGAAAGTAGTGATTT 889
QY 307 SerArgGlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHisGlyArgTrpGln 326
DB 890 TCGAGAGCGCTTCGATGTATCGAGAGTTACAGGCATCACCACATGGAAGATGGCAA 949
QY 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGln 346
DB 950 GCAAGATCGGCGGAGTCGCCGAAACAAAGACCTCTACTTGGGAACTTTTAGCACTGAG 1009
QY 347 GluGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaVal 366
DB 1010 GAAGAGAGCAGCAGAGCTTACGATATAGCTGCATTAAGTTTAGAGGACTTAATCGAGTG 1069
QY 367 ThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuPro 386
DB 1070 ACCAATTCGAGATCAACCGGTACGAGTGAAGGCCANTTCTAGAGAGTAGCACTCTTCCC 1129
QY 387 Ile-----GlySerAlaAlaLysArgLeuLysGluAla----- 397
DB 1130 ATCGGAGGAGCGCAGCTAAACGGCTCAAGGCTCAAGAGCTCAAGCTTTGAGTCTTCAAGGAA 1189
QY 397----- 397
DB 1190 CGCGAGCGGAGATGATAGCCCTTGGTTCAAGTTTCCAGTACGCTGGTGGTGGCTCGAGACA 1249
QY 398-----AsnArg 399
DB 1250 GGCTCTGGCTCCACCTCATCAAGACTTCAGCTTCAACCTTACCCTTAAGCATTTCAACAA 1309
QY 400 ProValProSerMetMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSer 419

```

```

Db 1310 COATTAGACGCTTTCTATCTCTTCAGAACATGACATCTCTCAATTAACAACAACAAT 1369
QY 420 GlyTyrGlnAsnAlaValGlnHisGlnGlyValAspLeuSerLeu---LeuHis 438
Db 1370 GCTCAGATCTCTCTCTTTAATCACCATAGCTATATCCAGACAACTTCATCTCCAC 1429
QY 439 GlnHis-----GlnGluArgTyrAsn 445
Db 1430 CAACAGACCAACAATTACTTTCAGCAACAGTCGAGCCAGAACTCTCAGCAGCTTCAAT 1489
QY 446 GlyTyrTyrTyr----- 449
Db 1490 GGTATCTTCATAGCAATCCGCTCTGCTTCATGACTTGCTCTACCTCTATCGTTGAC 1549
QY 450 -----AsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCysPheLysGln 465
Db 1550 AACAAATAATAACAATGGAGGCTCTAGTGGAGCTACAACACTGCAGCA----- 1597
QY 466 GluAspAspGlnHisPheLeuSerAsn----- 475
Db 1598 -----TTCTTTGGGAACCAACGGTATGTTGTTGGTCCAGCTCGACT 1639
QY 476 -----ThrGlnSerLeuMetThr---AsnIleAspHisGlnSerSerValSerAsp 491
Db 1640 GTTGATCGACGAGGAGCTTCCAAACCGTTAAACAGATTACGATATGCCTTCCAGTGAT 1699
QY 492 AspSerValThrValCysGlyAsnValValGlyTyrGlyGlyTyrGlnGlyPheAlaAla 511
Db 1700 -----GAAACCGAGGGTATAGTGTGG----- 1723
QY 512 ProValAsnCysAspAlaTyrAlaAlaSerGluPheAspTyrAsnAlaArgAsnHisTyr 531
Db 1723 ----- 1723
QY 532 TyrPheAlaGlnGlnGlnGlnThrGlnGlnSerProGlyGlyAspPheProAlaAlaMet 551
Db 1724 -----ACCAGTGAGTCTGTTCCAGGGGTCAAAACCT----- 1753
QY 552 ThrAsnAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGlyGluValAlaPro 571
Db 1754 -----GGTGGT----- 1759
QY 572 ThrPheThrValTyrAsnAsp 578
Db 1760 GTTTTCACTATGTGAATGAG 1780

```

RESULT 8

```

US-10-225-066A-855
; Sequence 855, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROWN, Pierre E
; TITLE OF INVENTION: Field-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MB10036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 09/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049

```

```

; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 855
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-855

```

```

Alignment Scores:
Pred. No.: 6,26e-112 Length: 1818
Score: 1047.50 Matches: 262
Percent Similarity: 46.96% Conservative: 70
Best Local Similarity: 37.06% Mismatches: 110
Query Match: 33.58% Indels: 265
DB: 13 Gaps: 23

```

US-09-980-364-2 (1-579) x US-10-225-066A-855 (1-1818)

```

QY 2 AsnAsnAsnTrpLeuGlyPhe-----SerLeuSerProTyrGlu 14
Db 65 TCTAACAACTGGCTTGGCTTCTCTTTCACGGAACAACCTCTCTTCCTCCATCAA 124
QY 15 GlnAsnHisHisArgLysAspValTyrSerSerThrThrThrValValAspValAla 34
Db 125 TACAAC----- 130
QY 35 GlyGluTyrCysTyrAspProThrAlaAlaSerAspGluSerSerAlaIleClnThrSer 54
Db 130 ----- 130
QY 55 PheProSerProPheGlyValValValAspAlaPheThrArgAspAsnAsnSerHisSer 74
Db 131 -----CTTGGCTTGGTCAGCAGCATATG-----GACAACCTTTTCAACA 172
QY 75 ArgAspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGly 94
Db 173 CAAGAGTGAATATG-----ATCAATCCACACACGGTGGAGGAGGAGATGAA 217
QY 95 -----ProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThr 110
Db 218 GGAGGAGAGGTTCGCGGCTTCTCGGTGAGCAACCGGAGCAAAACCA 277
QY 111 AsnGluAsnValGlyAspGlySerGlyCysTyrGlyGlyGlyAspGlyGlyGly 130
Db 278 TCCAAACCACTAGTAGCTTACAACGAGCTCAGACTACTACTTCCATACCAAT----- 328
QY 131 GlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProVal----- 147
Db 329 -----AGCTTGCTAGCTGCTCCAAATCAAAACGATGCTGTTGTAGCAGCT 373
QY 148 -----AspAsnValAspAsn-----GlnGluAsnGlyAsnAla 158
Db 374 TGTGACTCCAATACTCTTAACAACAGTAGCTATCATGAGCTTCAAGAGAGTGTCACAAT 433
QY 159 AlalysGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSer 178
Db 434 CTACAGTCACCTACTTGTTCATGGGACCAACC-----GCT 469
QY 179 AsnAsnAsnValVal-----AlaGlnGlyLysThrIleAspAspSerValGlu--- 194
Db 470 GGTAATATGTTGTAGACAAAGCTTCCACTCCAGACCAACCGGGGATAACGCTAGCGGT 529
QY 195 -----AlaThrProLysLysThrIleGluSerPheGlyGln 206
Db 530 GGAGCACTAGCCGTTGTTGAGACCGCCACGCAACGACGTCATTTGGACACTTTTCGGA 589
QY 207 ArgThrSerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHis 226

```


Db 65 TCTAACAACTGGCTGGCTTCTCTTTTCCCGAACTCTTTCTTCTTCTCTCTCATGAA 124
 QY 15 GlnAsnHisHisArgIysAspValTyrSerSerThrThrThrValValAspValAla 34
 Db 125 TACAAC----- 130
 QY 35 GlyGluTyrCysTyrAspProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSer 54
 Db 130----- 130
 QY 55 PheProSerProPheGlyValValAspAlaPheThrArgAspAsnAsnSerHisSer 74
 Db 131-----CTTGGCTTGTTCAGCACCATATG-----GACACCTTTTCAACAA 172
 QY 75 ArgAspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGly 94
 Db 173 CAAGAGTGGATATG-----ATCAATCCACACCGTGGAGGAGGAGATCAA 217
 QY 95-----ProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThr 110
 Db 218 GGAGGAGAGGTTCCAAAGTGGCCGATTTCTCGTGTGACAAACCGGACGAAACCAA 277
 QY 111 AsnGluAsnValGlyAspGlySerGlyCysTyrGlyGlyGlyAspGlyGlyGly 130
 Db 278 TCCAAACCACCTAGTAGTTACACAGCACTCAGACTACTTCCATACCAAT----- 328
 QY 131 GlySerLeuGlyLeuSerMetIleIysThrTrpLeuArgAsnGlnProVal----- 147
 Db 329-----AGTTTGATGCTAGCTGCTCAATCAACGATGCTGTTGTAGCAGCT 373
 QY 148-----AspAsnValAspAsn-----GlnGluAsnGlyAsnAla 158
 Db 374 TGTGACTCCAATCTCTCAACAGTAGCTATCATGAGTTCAGAGAGTGTCTCAAT 433
 QY 159 AlaySerGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSer 178
 Db 434 CTACAGTCACCTTACTTTGTCCATGGGACCACC-----GCT 469
 QY 179 AsnAsnValVal-----AlaGlnGlyLysThrIleAspAspSerValGlu--- 194
 Db 470 GGTAAATATGTTAGACAAGTTCACCATCGAGACCACCGGGGATAACGCTAGCGGT 529
 QY 195-----AlaThrProLysLysThrIleGluSerPheGlyGln 206
 Db 530 GGAGCACTAGCCGTGTTTGACGCGCCACGCAAGACGTCATTTGGACACTTTCCGACAA 589
 QY 207 ArgThrSerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHis 226
 Db 590 CGAACCTCGATCTATCTGTGTGTGCACAAAGATCGATGGACTGCTCGATATGAGGCTCAT 649
 QY 227 LeuTrpAspAsnSerCysLysArgGlyGlnThrArgLysGlyArgGlnValTyrLeu 246
 Db 650 CTATGGCAATAATAGTTAGAGGAAGGAGCCAGTCTAGGAAGAGAGCAAGTTTACTTG 709
 QY 247 GlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaIleLeuLysTyr 266
 Db 710 GTTGATATGACAAAGAAAGATGAGCAAGATCATATGATCTAGTCGACCTTAAGTAC 769
 QY 267 TrpGlyThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGlu 286
 Db 770 TGGGGTCTCTCACTACTACTATTTCCCATTAACAACTACGAGAAAGAGTAGAGGAA 829
 QY 287 MetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPhe 306
 Db 830 ATGAGCACATGACGACAGCAAGATGCTGGCTGCGCATTAGAGGAAAGATAGTGATTT 889
 QY 307 SerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHisGlyArgTrpGln 326
 Db 890 TCGAGAGGCCCTCTCGATGATCGAGAGATACAAAGGCATCCCAACATCGAAGATGGCAA 949
 QY 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGln 346
 Db 950 GCAAGGATCGGCCGAGTGGCGGAAACAAAGACCTCTACTTGGAACTTTTACGACTGAG 1009

QY 347 GluGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaVal 366
 Db 1010 GAAGAAGCAGCAGCAAGCTTACGATATAGCTGCAATAAAGTTTAGAGGACTTAATCGAGTG 1069
 QY 367 ThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuPro 386
 Db 1070 ACCAACTTCGAGATCAACCGGTACGAGTGAAGCAATTTCTAGAGAGTAGCCTCTTCCC 1129
 QY 387 Ile--GlySerAlaAlaLysArgLeuLysGluAla----- 397
 Db 1130 ATCGGAGGAGGCGCAGCTAAACGGCTCAAGAAAGTCAAGCTCTTTCAGTCTTCAAGGAAA 1189
 QY 397----- 397
 Db 1190 CGCGAGGCGAGATGATAGCCCTTGGTTCAAGTTTCCAGTACGGTGGCTCGAGCACA 1249
 QY 398-----AsnArg 399
 Db 1250 GGCTCTGGCTCCACCTCATCAAGACTTCAGCTTCAACCTTACCTCTTAAGCATTTCAACAA 1309
 QY 400 ProValProSerMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSer 419
 Db 1310 CCATTAGAGCCTTTTCTATCTCTTCAGAACATGACATCTCTCATTACAAACAAAT 1369
 QY 420 GlyTrpGlnAsnAlaAlaValGlnHisGlnGlyValAspLeuSerLeu---LeuHis 438
 Db 1370 GCTCAGCATTCCTCTCTTTAATCACCATAGTATATCCAGACACAACCTTCATCTCCAC 1429
 QY 439 GlnHis-----GlnGluArgTyrAsn 445
 Db 1430 CAAAGACCAACAANTACTTCGAGCAACAGTCGAGCCAGAACTCTCAGCAGCTTACAAT 1489
 QY 446 GlyTyrTyrTyr----- 449
 Db 1490 GCGTATCTTCATAGCAATCCGCTCTGCTTCATGGACTGTCTCTACCTCTATCGTTGAC 1549
 QY 450-----AsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCysPheLysGln 465
 Db 1550 AACATATATACATGAGGAGCTCTAGTGGAGCTACACACTGCAGCA----- 1597
 QY 466 GluAspAspGlnHisHisPheLeuSerAsn----- 475
 Db 1598-----TTTCTTGGAAACCAACCGTATTTGGTATTTGGTCCGCTCGACT 1639
 QY 476-----ThrGlnSerLeuMetThr---AsnIleAspHisGlnSerSerValSerAsp 491
 Db 1640 GTTGATCGACCGGAGGAGTTTCCAAACCGTTAAACAGATACGATATGCTTCCAGTGAT 1699
 QY 492 AspSerValThrValCysGlyAsnValValGlyTyrGlyGlyTyrGlnGlyPheAlaAla 511
 Db 1700-----GGAACCGGAGGTATAGTGTGG----- 1723
 QY 512 ProValAsnCysAspAlaTyrAlaAlaSerGluPheAspTyrAsnAlaArgAsnHisTyr 531
 Db 1723----- 1723
 QY 532 TyrPheAlaGlnGlnGlnThrGlnSerProGlyGlyAspPheProAlaAlaMet 551
 Db 1724-----ACCAGTGAGTCTGTTCCAGGGGTCAACCCCT----- 1753
 QY 552 ThrAsnAsnValGlySerAsnMetTyrTyrHisGlyGluGlyGlyGluValAlaPro 571
 Db 1754-----GGTGGT----- 1759
 QY 572 ThrPheThrValTrpAsnAsp 578
 Db 1760 GTTTTCACTATGTGGAATGAG 1780

RESULT 10

US-10-374-780A-333

; Sequence 333, Application US/10374780A

; Publication No. US20040019927A1

GENERAL INFORMATION:
 APPLICANT: Sherman, Bradley K
 APPLICANT: Riechmann, Jose Luis
 APPLICANT: Jiang, Cai-Zhong
 APPLICANT: Heard, Jacqueline E
 APPLICANT: Heake, Volker
 APPLICANT: Creelman, Robert A
 APPLICANT: Ratcliffe, Oliver
 APPLICANT: Adam, Luc J
 APPLICANT: Reuber, T. Lynne
 APPLICANT: Keddle, James
 APPLICANT: Bryon, Pierre E
 APPLICANT: Pilgrim, Marsha L
 APPLICANT: Dubell III, Arnold T
 APPLICANT: Pineda, Omaira
 APPLICANT: Yu, Guo-Liang
 TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
 FILE REFERENCE: MBI-0047 CIP
 CURRENT APPLICATION NUMBER: US/10/374,780A
 CURRENT FILING DATE: 2003-02-25
 PRIOR APPLICATION NUMBER: 09/837,944
 PRIOR FILING DATE: 2001-04-18
 PRIOR APPLICATION NUMBER: 60/310,847
 PRIOR FILING DATE: 2001-08-09
 PRIOR APPLICATION NUMBER: 09/934,455
 PRIOR FILING DATE: 2001-08-22
 PRIOR APPLICATION NUMBER: 60/336,049
 PRIOR FILING DATE: 2001-11-19
 PRIOR APPLICATION NUMBER: 60/338,692
 PRIOR FILING DATE: 2001-12-11
 PRIOR APPLICATION NUMBER: 10/171,468
 PRIOR FILING DATE: 2002-06-14
 PRIOR APPLICATION NUMBER: 10/225,066
 PRIOR FILING DATE: 2002-08-09
 PRIOR APPLICATION NUMBER: 10/225,067
 PRIOR FILING DATE: 2002-08-09
 PRIOR APPLICATION NUMBER: 10/225,068
 PRIOR FILING DATE: 2002-08-09
 NUMBER OF SEQ ID NOS: 2906
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 333
 LENGTH: 1818
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 FEATURE:
 OTHER INFORMATION: G1793
 US-10-374-780A-333

Alignment Scores:
 Pred. No.: 6.26e-112 Length: 1818
 Score: 1047.50 Matches: 262
 Percent Similarity: 46.96% Conservative: 70
 Best Local Similarity: 37.06% Mismatches: 110
 Query Match: 33.58% Indels: 265
 DB: 16 Gaps: 23

US-09-980-364-2 (1-579) x US-10-374-780A-333 (1-1818)
 QY 2 AsnAsnAsnTrpLeuGlyPhe-----SerLeuSerProTyrGlu 14
 DB 65 TCTAACAACTGGCTGGCTTTCCTCTTCCAGCAACACTCTCTTGGCTCTCTCATGAA 124
 QY 15 GlnAsnHisHisArgLysAspValTyrSerSerThrThrThrValValAspValala 34
 DB 125 TACAAAC-----130
 QY 35 GlyGluTyrCysTyrAspProThrAlaAlaSerAspGluSerSerAlaIleGlnThrSer 54
 DB 130-----130
 QY 55 PheProSerProPheGlyValValValAspAlaPheThrArgAspAsnSerHisSer 74
 DB 131-----CTTGGCTTGGTCAGCGACCATATG-----GACAACTCTTTCAAACA 172

QY 75 ArgAspTrpAspIleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAspGly 94
 DB 173 CAAGAGTGGATATG-----ATCAATCCACACCGGTGGAGGAGGATGAA 217
 QY 95 -----ProLysLeuGluAsnPheLeuGlyArgThrThrThrIleTyrAsnThr 110
 DB 218 CGAGGAGAGGTTCAAAAGTGGCGGATTTCTCGGTGTGAGCAAAACCGACGAAACCAA 277
 QY 111 AsnGluAsnValGlyAspGlySerGlySerGlyCysTyrGlyGlyGlyAspGlyGlyGly 130
 DB 278 TCCAAACCACTAGTAGCTTCAACAGCACTCAGACTACTACTTCCATATCACAAT-----328
 QY 131 GlySerLeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProVal-----147
 DB 329 -----AGCTTGATGCCTAGCGTCCAAACAAACGATGCTGTTGTAGCAGCT 373
 QY 148 -----AspAsnValAspAsn-----GlnGluAsnGlyAsnAla 158
 DB 374 TGTGACTCCAATACTCCTAACACAGTAGCTATCATGAGCTTCAAGAGAGTCTCACAAT 433
 QY 159 AlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCysAspAsnAsnAspSer 178
 DB 434 CTACAGTCACTACTTTTGTCCATGGGACCACC-----GCT 469
 QY 179 AsnAsnAsnValVal-----AlaGlnGlyLysThrIleAspAspSerValGlu---194
 DB 470 GGTATATATGTTGTAGACAAAGCTTCAACATCCGAGACACCGGGGATAACGTAGCGGT 529
 QY 195 -----AlaThrProLysLysThrIleGluSerPheGlyGln 206
 DB 530 GGAGCACTAGCGGTGTTGAGACGGCCACGCGCAGAGCTGCATCTGGACACTTCGGACAA 589
 QY 207 ArgThrSerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHis 226
 DB 590 CGAACTCGAICTATCGTGTGTCAAGACATCGATGCTGCTCGATATGAGCTCAT 649
 QY 227 LeuTrpAspAsnSerCysLysArgGluGlnThrArgLysGlyArgGlnValTyrLeu 246
 DB 650 CTATGGGATAATAGTTGTAGAGGAGGCCAGTCTAGGAAAGGAGGACAACTTACTTGT 709
 QY 247 GlyGlyTyrAspLysGluGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyr 266
 DB 710 GGTGGATATGCAAAAGAGATAAAGCAGCAAGATCATATGATCTAGCTGCATTAAGTAC 769
 QY 267 TrpGlyThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGlu 286
 DB 770 TGGGTCTCTCAACTACTACTAATTTCCCAATTACAAACTACGAGAAAGAGTAGAGGAA 829
 QY 287 MetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPhe 306
 DB 830 ATGAACACATGACGAGACAGAGATTCGTGGCTGCAATAGAGGAAAGTATGATTT 889
 QY 307 SerArgGlyAlaSerIleTyrArgGlyValThrArgHisHisGlnHisGlyArgTrpGln 326
 DB 890 TCGAGAGCGCTTCGATGTATCGAGGAGTTACAAGGCATCACCAACATCGAAGATGGCAA 949
 QY 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGln 346
 DB 950 GCAAGGATCGCGCGATCGCCGAAACAAAGACCTCTACTTGGGAACCTTTTAGCTAGT 1009
 QY 347 GluGluAlaAlaGluAlaTyrAspIleAlaAlaIleLysPheArgGlyLeuThrAlaVal 366
 DB 1010 GAAGAAGCAGCAGAGACTTACGATATAGCTGCAATAAAGTTTAGAGGACTTAATGCACTG 1069
 QY 367 ThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuPro 386
 DB 1070 ACCAACTTCGAGATCAACCGGTACGCGTGAAGGCCATTTCTAGAGTAGTACACTCTTCCC 1129
 QY 387 Ile---GlySerAlaAlaLysArgLeuLysGluAla-----397
 DB 1130 ATCGAGGAGCGCGCACTAAACGGCTCAAGAGAGCTCAAGCTCTTGTAGTCTTCAGAGAA 1189

QY 232 CysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeuGlyGlyTyrAspLys 251
Db 597 TGCAGAGAGAGAGGCGAAAGTAGGAAGAGGCGAAGTTATTTCGGTGGCTATACACAG 656
QY 252 GluGluValAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyrTyrGlyThrThr 271
Db 657 GAAGATTAAGCTGCAAGAGCTTATGATCTTCGGCTCTCAAGTACTGGGTCGACCACT 716
QY 272 ThrThrAspPheProMetSerGluTyrGluLysGluValGluGluMetLysHisMetThr 291
Db 717 ACTACCAACTTTCGGATTTCCTCAACTATGAGAGAGGAGCTCGAAGAGATGAAGACATGAT 776
QY 292 ArgGlnGluTyrValAlaSerLeuArgArgLysSerSerGlyPheSerArgGlyAlaSer 311
Db 777 AGCAAGAGTTCGTTCTCTCCCGAGAGAAAGTAGTGATTTCTAGGGTGGCTTCA 836
QY 312 IleTyrArgGlyValThrArgHisHisGlnHisGlyArgTyrGlnAlaArgLysGlyArg 331
Db 837 ATTACAGAGGAGTGAACAGGCATCATCAACATGGTAGGTGGCAAGCAAGAATTTGAAGA 896
QY 332 ValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGlnGluAlaAlaGlu 351
Db 897 GTTCAGGCAACAAAGATCTCTATCTTGCCACATTTAGCACCCCAAGAGAGAGCTGAA 956
QY 352 AlaTyrAspIleAlaAlaLysPheArgGlyLeuThrAlaValThrAsnPheAspMet 371
Db 957 GCCTATGATATTCAGCAATCAAGTTTAGAGGCTTAAACGGGTTACAAATTTTGATAG 1016
QY 372 AsnArgTyrAsnValLysAlaLeuLeuGluSerProSerLeuProIleGlySerAlaAla 391
Db 1017 AGCGGTACCATGATAACAGCATTCGCAACAGC---AATCTTCCCATTTGGAGCACTGAAC 1073
QY 392 LysArgLeuLysGluAlaAsnArgProValProSerMetMetMetLeuSerAsnVal 411
Db 1074 AACCAATCCAAACTGCAATGATTCACAT---TCTGATAGGAGC 1115
QY 412 SerGluSerGluAsnSerAlaSerGly-----TrpGlnAsnAlaAla 425
Db 1116 ACCGATGAGGAGGAGATCATCTCTGGATCATCGGCGATGACCTTATCATCTCAGCCCGCT 1175
QY 426 ValGlnHisGlnGlyValAspLeuSerLeuHisGlnHisGlnGluArgTyrAsn 445
Db 1176 GCAAGCTCCAGTAGGCTTTGCAATTAACATCAAGCAACCCATCCGATATTGGTCA 1235
QY 446 -----GlyTyrTyrTyrAsnGlyGlyAsnLeuSerSerGluSerAlaArgAlaCys 462
Db 1236 AACGTTCTCGGATACATGATGATCTGCTGCTTCTTAAGCAGT---GCCAAGAACCCCTATA 1292
QY 463 PheLysGlnGluAspAspGlnHisHisPheLeuSerAsnThrGlnSerLeuMetThrAsn 482
Db 1293 TTTTCATCACTCACCAATGGATCGGTTTCCAAAGT---CCAAAGTGGATTCGTACCGGC 1349
QY 483 IleAspHisGlnSerSerValSerAspSerValThrValCysGlyAsnValValGly 502
Db 1350 AATCGTGGGATTCGCGATCAACGGA-----AGCATTAACATGGGTTATTCAAT 1400
QY 503 TyrGlyGlyTyrGlnGly----- 508
Db 1401 CGAGGTGGTATGTTGGGCGAGGAGCAACAGAGAGCAGTGTCTACTAATATCCCAACCCA 1460
QY 509 PheAlaAlaProVal-----AsnCysAspAlaTyrAlaAlaSerGluPheAspTyr 525
Db 1461 TTGGCAACGCCAATGCTTAAATTAATAGCGATTAATTAGAGCTGTCTCTCGGCTAT 1520
QY 526 AsnAlaArgAsnHisTyrTyrPheAlaGlnGln-----GlnGlnThrGlnGlnSerProGly 544
Db 1521 GGT-----GGTAACCTGGAAATCTCGACAGTCGTTTCAATCTACTTATCAAACTGCAAG 1574
QY 545 GlyAspPheProAlaAlaMetThr----- 552
Db 1575 CCAAGTCTCTCAGTGTTCCAACACCAACCAATTTTGGCATGGAATGAATCAGCTTCACATG 1634

QY 553 ---AsnAsnValGlySerAsnMetTyrTyrHisGly-----GluGly 565
Db 1635 GACATTAACCTGTGGAATAAATCTG---TGGAGGCGCGATGCATTCAGAGTTGGAAGGG 1691
QY 566 Gly-----GlyGluValAlaProThr 572
Db 1692 GGTAGGTTGTTTAACCTTTCTTGAATAGGTGAGATTGGAATGACA 1736
RESULT 12
US-10-183-687-484
; Sequence 484, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Farnodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: BBI458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 484
; LENGTH: 1959
; TYPE: DNA
; ORGANISM: Catalpa sp.
US-10-183-687-484
Alignment Scores:
Pred. No.: 1,15e-106 Length: 1959
Score: 1003.00 Matches: 238
Percent Similarity: 54.20% Conservative: 46
Best Local Similarity: 45.42% Mismatches: 84
Query Match: 32.16% Indels: 156
DB: 13 Gaps: 17
US-09-980-364-2 (1-579) x US-10-183-687-484 (1-1959)
QY 2 AsnAsnAsnTyrLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLysAsp 21
Db 109 CATCAGAACTGCTCGCTTCTCGCTTTCC-----AACCATCT----- 147
QY 22 ValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAspPro 41
Db 148 -----CTGTTTCACCCC 159
QY 42 ThrAlaAlaSerAspGluSerSerSerAlaIleGlnThrSerPheProSerPheGlyVal 61
Db 160 TCGTTCAACTCTAATCCCTCCCTCCAGCTGC-----GGGTT 195
QY 62 ValValAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAspIleAsnGly 81
Db 196 GTG-----AATCAAGAAATGCGGT-----GCGGT 222
QY 82 CysAlaCysAsnAsnIleHisAsnAsp-----GluGlnAspGlyProLysLeuGluAsn 99
Db 223 GGTCCGCAACAGATCATCTCTCGGTTCTTTCAGTGGAGGCGCAAGCTGGAGAT 282
QY 100 PheLeuGlyArgThrThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerGly 119
Db -----

Score:	996.50	Matches:	252
Percent Similarity:	50.63%	Conservative:	70
Best Local Similarity:	39.62%	Mismatches:	182
Query Match:	31.95%	Indels:	132
DB:	13	Gaps:	20
US-09-980-364-2 (1-579) x US-10-183-687-490 (1-1954)			
Qy	7	GlyPheSerLeuSerProTyrGluGln-----AsnHisHisArgLysAspValTyrSer	24
Db	117	GGGGTCTCTCTCACTCCACGCTTCAGAGATTGATGAAGAGTTTCAGAGAGAAACCAAGAGC	176
Qy	25	SerThrThrThrValValAspValAlaGlyGluTyr:CysTyrAspProThrAlaAla	44
Db	177	GTGAGAGAGGTTACCATCATCTCTCTGTCACCCCTCACTTCTGTCTGTATGCCCCCTTCGC	236
Qy	45	SerAspGluSerSerAlaIleGlnThrSerPheProSerPheGlyValValAsp	64
Db	237	TCTGATGGCTCTCTCTGTGTGTCTGCTCTTT-----	269
Qy	65	AlaPheThrArgAspAsnAsnSerHisSerArgAspTyr-AspIle-----	79
Db	270	-----AGTCACCTCTGCTGCTGATCTCAAGAAATGGAGATATGATAGTCCAANTAGG	317
Qy	80	-AsnGlyCysAlaCysAsnIleHisAsnAspGluGlnAspGlyProLysLeuGluAs	99
Db	318	TGGAGGG-----AATTCCAATGAAGAAGGTCCTCAAGCTTGAAGA	356
Qy	99	nPheLeuGly-----	102
Db	357	CTTTTGGGTGTCTACTCAGACTCCCCCGCTAAGGTCITTTGCCAAGACTCTCAGCCCTGA	416
Qy	103	-----ArgThrThrThrIleTyrAsnThrAsnGluAsnValGlyAspGlySerG1	119
Db	417	CCAAAACCAAGCCAGAACAAATGTTCCAAATCAATGTAATGTG-----GCACC	467
Qy	119	ySerGlyCysTyr:GlyGly-----GlyAspGlyGlyGlyGlySerLeuG1	134
Db	468	AAGCTTTTGCCACCAAGATTCAGAGATTGAAACTGGGGATAATCTCAAAATCCTTCTTC	527
Qy	134	yLeu-----	136
Db	528	CTTGCTCCATCTTTCATGCATATTAATGAATCACTCAATGCTCTTATCCCCACCAATGG	587
Qy	136	rMetIleLysThrTrpLeuArgAsnGlnProValAspAsn-----ValAspAsnG1	153
Db	588	CATGTACAAGCTCTGTTGGCACAGACCCAGTCTTTCTCTGTATGAAACCTTCAAAATGA	647
Qy	153	nGluAsnGlyAsnAlaAlaLysGlyLeuSerLeuSerMetAsnSerSer-----	169
Db	648	GGCTAATGGGTGCAATTTCCAAATCTTTGTCCTCACTATGACCCCACTGTGCAAAATGG	707
Qy	170	-----Thr-SerCysAspAsnAsnAsnAspSerAsnAsnValValAlaG1	185
Db	708	GGTGGGTGCAATTCCTCTCTTCACTGATGAGGATAGTACGAAACGGGTCTATGGCGAA	767
Qy	185	nGlyLysThrIleAspAspSerValGluAlaThrProLysLysThrIleGluSerPheG1	205
Db	768	A-----TCTCATGTAGAGAACAGGTTCTCTCGCAAACTTATTTAGACATTTTGG	815
Qy	205	yGlnArgThrSerIleTyrArgGlyValThrArgHisArgTyrThrGlyArgTyrGluAl	225
Db	816	GCAAAGNACATCTCAATATCGTGTGTTTCAAGGCATAGATGGACTGGAAGATATGAGGC	875
Qy	225	aHisLeuTrpAspAsnSerCysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyr	245
Db	876	CCATTTCTGGGTAATAATAGTTGCAGAAAGGAAGGCGGCAAAACAGCAAGGAAGCAAGTTTA	935
Qy	245	rLeuGlyGlyTyrAspLysGluLysAlaAlaArgAlaTyrAspLeuAlaAlaLeuLys	265
Db	936	CTTGTGTGGTTATGATAGGAAGAAACGACAAAGGCTTACGACTTAGCTGCACTCAA	995
Qy	265	sTyrTrpGlyThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValG1	285

;; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
;; FILE REFERENCE: 38-21(53223)B
;; CURRENT APPLICATION NUMBER: US/10/424,599
;; CURRENT FILING DATE: 2003-04-28
;; NUMBER OF SEQ ID NOS: 285684
;; SEQ ID NO 100930
;; LENGTH: 2116
;; TYPE: DNA
;; ORGANISM: Glycine max
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT3847_62152C.1
US-10-424-599-100930

Alignment Scores:
Pred. No.: 8,54e-106 Length: 2116
Score: 996.00 Matches: 233
Percent Similarity: 61.59% Conservative: 62
Best Local Similarity: 48.64% Mismatches: 107
Query Match: 31.93% Indels: 77
DB: 13 Gaps: 14

US-09-980-364-2 (1-579) x US-10-424-599-100930 (1-2116)

QY 2 AsnAsnAsnTrpLeuGlyPheSerLeuSerProTyrGluGlnAsnHisArgLysAsp 21
DB 275 TCAGTAACCTGGATCGTCTCTCTCCCGCATGGAATGCTCGAACCCTCGAACCT 334
QY 22 ValTyrSerSerThrThrThrValValAspValAlaGlyGluTyrCysTyrAspPro 41
DB 335 CAGTTCGTCAA-----TACGAC--- 352
QY 42 ThrAlaSerAspGluSerSer-----AlaIleGlnThrSerPheProSerPro 58
DB 353 ---GCCGCTTCGACACTTCTCGCATCACTACTACGACAACTGTACACCAACGGG 409
QY 59 PheGlyValValValAlaAspAlaPheThrArgAspAsnAsnSerHisSerArgAspTrpAsp 78
DB 410 TGGGG-----AACGGAGCCTCAAGTTTGACGAGAAT 442
QY 79 IleAsnGlyCysAlaCysAsnAsnIleHisAsnAspGluGlnAsp-----Gly 94
DB 443 CTGAACACACGACGACGTGATCGTGTGAATCGTGTGCGACGAGCGTACGCCACCGCG 502
QY 95 ProLysLeuGlnAsnPheLeuGlyArgThrThrThrIleTyrAsnThrAsnGluAsn--- 113
DB 503 CCGAGCTGGAGGATTTCTCGCGACTCTCGCGTGTATGCGTACTCTCGACACCGCAG 562
QY 114 -----ValGlyAspGlySer----- 118
DB 563 ACGGAGACGACGACTCGTCTGACGACATCTACGACACACACACACACACACAC 622
QY 119 -----GlySerGlyCysTyrGlyGlyAspGlyGly-----GlySer 132
DB 623 CACCACACACACGCTTCTTCGCTACTTCGGGGTGACACACACACACATCTCAAGGCCAT 682
QY 133 LeuGlyLeuSerMetIleLysThrTrpLeuArgAsnGlnProValAspAsnValAsp-As 152
DB 683 ACTGATTCGAAGCTTTTCGACTTAACCTCGGCTCCGAGGTTGATTCGATCTGCAATC 742
QY 152 nGlnGluAsnGlyAsnAlaLysGlyLeuSerLeuSerMetAsnSerSerThrSerCy 172
DB 743 GGAAGAGCGGACGACGAGTTCGGGACTCACTCTATTGAGTCTCCCGCAACACGATTC 802
QY 172 sAspAsnAsnAspSer-AsnAsnAsnValValAlaGlnGlyLysThr---Ile-Asp 190
DB 803 GCCCGCTTCCTCGGTGGCACAACCGGAACCTTGTCTCGCTCGCGTGGACTGAGCTCCGAG 862
QY 191 AspSerValGluAla-----ThrProLysLysThrIleGluSerPheGlyGln 206
DB 863 AAGCCCGTTGTGGCGGCGAGTCCAAATAGTTCGAAGAAGATCGTGGATACCTTCGCGCAG 922
QY 207 ArgThrSerIleTyrArgGlyValThrArgHisArgTrpThrGlyArgTyrGluAlaHis 226

DB 923 CGGACTTCTATTACAGAGGTGTTACTAGGCACCCATGGACAGGAAGATATGAAGCGCAT 982
QY 227 LeuTrpAspAsnSerCysLysArgGluGlyGlnThrArgLysGlyArgGlnValTyrLeu 246
DB 983 CTATGGGACATAGTTTCAGAGAGGGGTCAAGCCAGAAAGGCGCTCAAGTTTATTG 1042
QY 247 GlyGlyTyrAspLysGluGlyAlaAlaArgAlaTyrAspLeuAlaAlaLeuLysTyr 266
DB 1043 GGTGATATGATAAGGAAGAAAGCCCGCAGAGCTTATGATTTGGCAGCTCTAAGATAC 1102
QY 267 TrpGlyThrThrThrThrAsnPheProMetSerGluTyrGluLysGluValGluGlu 286
DB 1103 TGGGTCCTCCACTGCTACCCCACTTCCCTGTTTCCCAATTATTCGAGGAGTGGAGAG 1162
QY 287 MetLysHisMetThrArgGlnGluTyrValAlaSerLeuArgLysSerSerGlyPhe 306
DB 1163 ATGAACATGTAAACAAGCAAGAATTTCGTCATCATTCGCGAGGAAAAGTAGTGGTTTC 1222
QY 307 SerArgGlyAlaSerIleTyrArgGlyValThrArgHisGlnHisGlyArgTrpGln 326
DB 1223 TCCAGGGAGCTTCCATATACAGAGGTGTAAAGGCATCATCAACAGGTAGGTGGCAA 1282
QY 327 AlaArgIleGlyArgValAlaGlyAsnLysAspLeuTyrLeuGlyThrPheGlyThrGln 346
DB 1283 GCAAGAAATGGCGGTGTAGCTGGAACAAGATTATATCTTGGGAACATTGCAACCGAG 1342
QY 347 GluGluAlaAlaGluAlaTyrAspIleAlaAlaLysPheArgGlyLeuThrAlaVal 366
DB 1343 GAGGAAGCAGCAGAGGCATATGATATTCAGCCATAAAGTTTCAGAGGTGCAACCGGTA 1402
QY 367 ThrAsnPheAspMetAsnArgTyrAsnValLysAlaIleLeuGluSerProSerLeuPro 386
DB 1403 ACCAATTTGAGATGATAGATATGATGTGGAGCTATATGAAGAGT---TCTCTTCCA 1459
QY 387 IleGlySerAlaAlaLysArgLeuLys-----GluAlaAsnArgProValProSer 403
DB 1460 GTGGGTGGGCGCAGCAAAACGCTTGAGGCTTCCCTTGAATCAGACGACAGAAAGCTCCT 1519
QY 404 MetMetMetIleSerAsnAsnValSerGluSerGluAsnSerAlaSerGlyTrpGlnAsn 423
DB 1520 GTG-----AACAGCAGCAGTCAGCAGAGAT 1546
QY 424 AlaAlaValGlnHisHisGlnGly---ValAspLeuSerLeuLeuHisGln 439
DB 1547 CCACAGTGTGGTAACGTGAGTGGTAGCATCAATTTCTCAGCCATTCATCAG 1597

Search completed: July 19, 2004, 20:17:48
Job time : 809 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 09:59:14 ; Search time 20 Seconds
(without alignments)
2784.745 Million cell updates/sec

Title: US-09-980-364-2
Perfect score: 3119
Sequence: 1 MNNWLGFLSPYEQNHHRK.....YVHGEGGGEVAFVFWNDN 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2475	79.4	581	2 T51580	ovule development
2	1040.5	33.4	516	2 P96549	hypothetical prote
3	997	32.0	566	2 T49988	ovule development
4	930	29.8	485	2 T03638	hypothetical prote
5	926	29.7	425	2 B96750	hypothetical prote
6	891	28.6	555	2 S71365	ovule development
7	654.5	21.0	332	2 D86295	hypothetical prote
8	619.5	19.9	308	2 H96827	protein F20B17.12
9	572	18.3	485	2 S54116	hypothetical prote
10	511	16.4	205	2 T47591	aintegumenta-like
11	475.5	15.2	433	2 T01574	indeterminate spik
12	465.5	14.9	464	2 C84686	probable AP2 domai
13	461.5	14.8	432	2 A85436	APETALA2 protein
14	459	14.7	436	2 B84845	probable AP2 domai
15	446	14.3	446	2 T03981	APETALA2-like prot
16	270	8.7	247	2 T06721	hypothetical prote
17	269	8.6	236	2 T02577	probable AP2 domai
18	197	6.3	292	2 T09030	hypothetical prote
19	180	5.8	358	2 D96579	hypothetical prote
20	162	5.2	314	2 E86452	protein FSU5.5 [im
21	157.5	5.0	277	2 T03927	AP2 domain transcr
22	155.5	5.0	336	2 B84594	transcription fact
23	153.5	4.9	207	2 T49897	probable AP2 domai
24	150	4.8	171	2 T00432	hypothetical prote
25	150	4.8	245	2 B86168	hypothetical prote
26	149.5	4.8	328	2 G86283	hypothetical prote
27	149	4.8	204	2 B86410	protein F3M18.20
28	148.5	4.8	261	2 B86354	hypothetical prote
29	148	4.7	245	2 B86407	hypothetical prote

ALIGNMENTS

RESULT 1

T51580

ovule development protein aintegumenta-like protein - Arabidopsis thaliana
N:Alternate names: protein T10B6_90

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000

C:Accession: T51580

R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asanizu, E.; Kotani, H.; Tabata, S.; Mew

Submitted to the Protein Sequence Database, August 2000

A:Reference number: Z25394

A:Accession: T51580

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-581 <SAT>

A:Cross-references: EMBL:AL391142

A:Experimental source: cultivar Columbia; BAC clone T10B6

C:Genetics:

A:Map position: 5

A:Introns: 78/1; 216/2; 244/1; 273/3; 298/2; 315/2; 341/1

A>Note: T10B6_90

Query Match 79.4%; Score 2475; DB 2; Length 581;

Best Local Similarity 81.3%; Pred. No. 2.3e-166;

Matches 474; Conservative 36; Mismatches 59; Indels 14; Gaps 7;

QY	3	NNWLGFLSPYEQNHHRKDVYSSITTTVDVAGEYCYDPTAASDESSAIOQSFSPFGV 62
DB	5	NNWLGFLSPHDQNHRTDVSSTRTAVDVAGGYCFDLAAPSDESSAVQTSFLSPFGV 64
QY	63	VDATRDNNSHSRDWDINGACNNIHNDQDQPKLENFLGRITTTTYNTNENVDGSGGCG 122
DB	65	LEAFTRDNNSHSRDWDINGACNNINNEQNGPKLENFLGRITTTTYNTNETVVDGNGD-C 123
QY	123	YGGGPGGGGSLGLSMIKTWLRNQPVDVNDQNGNAAGLSLSMNSSTSCDNNNDSNNV 182
DB	124	GGDGGGGGSLGLSMIKTWSNVANHQNGARGLSLSMNSSTSCDNNNNDND 182
QY	183	VAQGTIDDSVEATPKKTIESFGQTSYRGVTRHRTGRVEAHLWDNSCKREGQTRKR 242
DB	183	VVOENTVDVWETTPKKTIESFGQTSYRGVTRHRTGRVEAHLWDNSCKREGQTRKR 242
QY	243	QYVLGGYDKEEKAAYDLAALKYWGTTTTNFPMSYEKEVEEMKHTROYVASLRK 302
DB	243	Q---GGYDKEEKAAYDLAALKYWGTTTTNFPUSEYEKEVEEMKHTROYVASLRK 299
QY	303	SSGFGSGASIRGVTRHHQHWQARIQGVAGNKDLYLGTGTQBEAAEAYDIAAKFRG 362
DB	300	SSGFGSGASIRGVTRHHQHWQARIQGVAGNKDLYLGTGTQBEAAEAYDIAAKFRG 359
QY	363	LTAIVNFDNVRNVVAILESPLPGISAKRLKEANRPVPSMMISNNVSSSENSASGWQ 422
DB	360	LSAVTNFDNVRNVVAILESPLPGISAKRLKLVNVPVA-MVTSNNVSSSANNVSGWQ 418

QY 423 NAAVQHGGVDSLIIHQHRYNGYVYNGNLSSESARACFKQEDDQHHFLSNTQSLMTN 482
 Db 419 NTAQHHGGMDLSLLQQQERYVG-YVYNGNLSSTESTRVCFKQEEQHHFLNSPFWTN 477
 QY 483 IDHSSVSDSDSVTCGVVGGVGGGGAAP-----VNCDAVAAEFYDYNARHHYFA--Q 535
 Db 478 VDHSSSTSDSDSVTCGVVGGVGGGGAAP-----VNCDAVAAEFYDYNARHHYFA--Q 537
 QY 536 QOQOQSGGDFPAAMTNVGNMNYHGGGGEVAPFTVWND 578
 Db 538 QOQOQSGGDFPAAMTNVGNMNYHGGGGEVAPFTVWND 580

RESULT 2

F96549
 N;Alternate names: protein F11M15.6 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: F96549
 R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; PMID:21016719; PMID:11130712
 A;Accession: F96549
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-516 <STO>
 A;Cross-references: GB:AB005173; NID:94836931; PIDN:AAD30633.1; GSPDB:GN00141
 C;Genetics:
 A;Gene: F11M15.6
 A;Map position: 1

Query Match 33.4%; Score 1040.5; DB 2; Length 516;
 Best Local Similarity 44.6%; Pred. No. 1.7e-65;
 Matches 254; Conservative 70; Mismatches 114; Indels 131; Gaps 22;
 QY 86 NIHDEQDG---PKLENFLGRITTT-IYNTNEN-----VGDGSGGCGYGGDGGGSLGL 135
 Db 3 NPHGGGEGGEPKVPADFLGVSKGDHHTDHLVYNDIHQTNADSYFQTN---SL-L 57
 QY 136 SWIKTWLRNQPVDVNDQENGNAAKGLSLMNS-----TSCDNNDSN 179
 Db 58 PTVTTCASNAP-NNYELOESAHNLQSLTSLMSTGCAAAAEVATYKASPAETSADNSSTT 116
 QY 180 NNVAQGTIDDSVEATPKTIESFGORTSYRGVTRHRTGTYEAHLWNSCKREGQTR 239
 Db 117 N---TSGGAI---VEATPRITLETFGRTSYRGVTRHRTGTYEAHLWNSCKREGQSR 170
 QY 240 KGRQYLGYYDKEEKAARAYDLAALKYWGTTTITNPFMSYEKEVEEMKMTQEVVASL 299
 Db 171 KGRQ---GGYDKEEKAARAYDLAALKYWGTPSTTTNPFITNYEKEVEEMKMTQEVVASI 227
 QY 300 RKSGFSGASIVRGVTRHHQHWQARIGRVAGNKDLYLGTGTOEEAAEAYDIAAIK 359
 Db 228 RKSGFSGASMRVGRVTRHHQHWQARIGRVAGNKDLYLGTGTPSTEEAAEAYDIAAIK 287
 QY 360 FRGLTAVTNFDMRYNKAILESPLI-GSAAKRLKEA-----NRPVPMYISNNVSE 413
 Db 288 FRGLNAVTFEINRYDVKAILLESNTLPGGGAARLKEAQALESRRKREMIALGSNFHQ 347
 QY 414 SENSASQWNA-----VOHQGV-----DLS----- 435
 Db 348 -YGAASGSSVASSRRLQLPYPLSIQOPFHLHHQPLLTQLQNNNDISQVHDSFSYIQT 406
 QY 436 LLHQHRYNGYVYNGNLSSESARACFKQEDDQHHFLSNTQSLMTNIDHQSSVSDSDSVT 495

Db 407 QLHLHQOQTNNYLOSSHTSQL-----YNAYLQSNFGLL-----HGFVSDNN-- 448
 QY 496 VCGNVVGYGGYGAAPVNCDAVAAEFYDYNARHHYFAQQOQTOQSPGDFPAAATNNV 555
 Db 449 ---NTSGFLGNGGIGGSSSTVGSAAEFEPFAVKDY-----DMPPS----- 487
 QY 556 GSNMYHGGGGEVAP-----TFTVWND 578
 Db 488 GGATGCGGNWNGESAQGSNPGGVFTWNE 516

RESULT 3

T49988
 N;Alternate names: protein F12B17.140
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C;Accession: T49988
 R;Bevan, M.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, April 2000
 A;Reference number: Z25026
 A;Accession: T49988
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-566 <BEV>
 A;Cross-references: EMBL:AL333995; GSPDB:GN00063; ATSP:F12B17.140
 A;Experimental source: cultivar Columbia; BAC clone F12B17
 C;Genetics:
 A;Gene: ATSP:F12B17.140
 A;Map position: 5
 A;Introns: 33/1; 259/2; 287/1; 316/3; 341/2; 384/1

Query Match 32.0%; Score 997; DB 2; Length 566;
 Best Local Similarity 40.4%; Pred. No. 2.2e-62;
 Matches 247; Conservative 58; Mismatches 127; Indels 180; Gaps 18;
 QY 43 AASDESSAIQTSFSPFGVVDVDAFTRDNNSHSRDWDINGCACNNIHNDQDQPKLENFLG 102
 Db 52 AASMDSTILITTFVDP-----QSHHSQNHIPKLEDFLG 84
 QY 103 RT-----TTTY-----NTNEN 113
 Db 85 DSSSIVRYSDNSQTDQDSSLTQIYDPRHHNQTFYSDHDFKTMAGFQSAFSTNSGE 144
 QY 114 VGDGSGG-----CYGGDGGGSLGLSMIKTWLRNQPVDVNDQ 153
 Db 145 VDDASIGRTHLAGDYLGHVVESSGPELGFHGGSTGALSLGVNWN--TNHRDN-DNH 201
 QY 154 ENGNAAGLSLSMNSSTSCDNNDSNNNVVAGK-----TIDDSVEATPKTIESFGORTS 209
 Db 202 YRGN-----NNGERINNNDNEKTDSEKEKAVVAVETSDCSNKKIADTFGORTS 252
 QY 210 IYRGVTRHRTGTYEAHLWNSCKREGQTRGROVYLGYYDKEEKAARAYDLAALKYWG 269
 Db 253 IYRGVTRHRTGTYEAHLWNSCKREGQTRGROVYLGYYDKEEKAARAYDLAALKYWNA 309
 QY 270 TTTTNFPMSEYEKEVEEMKMTQEVVASLRRKSGFSGASIVRGVTRHHQHWQARI 329
 Db 310 TATTNFTNYSYEKEVEEMKMTQEVVASLRRKSGFSGASIVRGVTRHHQHWQARI 369
 QY 330 GRVAGNKDLYLGTGTOEEAAEAYDIAIKFRGLTAVTNFDMRYNKAILESPLTGS 389
 Db 370 GRVAGNKDLYLGTGTOEEAAEAYDIAIKFRGLNAVTFEINRYDVKAILLESNTLP 428
 QY 390 AAKRLKEANRPVPMYISNNVSESENSASQWNAVCHQGVDSLHLHQHRYNGY 449
 Db 429 AAKRLK-----LSLEAAASSEQKPFILGHQ-----LHFFQOQ----- 461
 QY 450 NCGNLSSESARACFKQEDDQHHFLSNTQSLMTNIDHQSSVSDSDSVTCGVVGGYQGF 509
 Db 462 -----QQOQLQSSPNHSSINFALCPNSAVOSQQIIPCG----- 496

A;Cross-references: EMBL:U40256; NID:gl209098; PID:gl209099
R;Bevan, M.; Van der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15484
A;Accession: T06022
A;Molecule type: DNA
A;Residues: 1-555 <BEV>
A;Cross-references: EMBL:AL035709; GSPDB:GN00062; ATSP:T28119.30
A;Experimental source: cultivar Columbia; BAC clone T28119
C;Genetics:
A;Gene: ATSP:T28119.30; ANT
A;Map position: 4
A;Introns: 111/1; 289/2; 317/1; 320/1; 349/3; 391/2; 417/1

Query Match 28.6%; Score 891; DB 2; Length 555;
Best Local Similarity 40.8%; Pred. No. 6.3e-55;
Matches 227; Conservative 56; Mismatches 122; Indels 152; Gaps 17;

QY 4 NWLGFSLSPYEQ----NHRKDVYSSTTTTVDVAGYCYDPTAASDESSA----IQTSF 55
DB 17 NLLGFSLSNNMKWGGCGRAIYSSSTS-----SATSSTSSVPPQLVVDN 63
QY 56 PSPFGV-----VVDAFTDNNNS-HSRD-----WDINGCA 83
DB 64 TSNFGVCGSNPNNGIYSHMSVPLRSDGSLCLMEALNRSHSHNHQDSSPKVEDFFGTH 123
QY 84 CNNIHNDQDGPKEFLGRITTYTNNVGD-----CGSGCYGGGDG---GGG 131
DB 124 HNTSHKEMDLSDLSLYNTTTPNTTNPQEFPSPTQENHEETRYGNPDLTHGG 183
QY 132 SLGLSMIKTLNRPQVNDVNDQNGNAAGLSLNN--SSTSC-----DNNDS 178
DB 184 SFNVGVY-----GEFQQLSLSMSPGSSSCITGSHHHQONQNHQS 226
QY 179 NNVVVAQCKTIDDSV-----BAPKKTIESFGORTSIYRG 213
DB 227 QNHQOISALVETSVGFETTTMAAKKRGQEDVVVQKQIVHRKSIDTFGQTSYRG 286
QY 214 VTRHRTGRYEAHLWNSCKREGQTRKGRQVYLGGYDKEEKAARAYDLAALKYWGTTTTT 273
DB 287 VTRHRTGRYEAHLWNSCKREGQTRKGRQVYLGGYDKEEKAARAYDLAALKYWGPTSTH 346
QY 274 NPPMSEYEVEEMKHTROYVAVSLRRKSSGFSRGASIVRGVTRHHQHGRQWQARIGRVA 333
DB 347 NFSAEYQKEIDMKNTROYVAVSLRRKSSGFSRGASIVRGVTRHHQHGRQWQARIGRVA 406
QY 334 GNKOLYLGTFGTQBEAAAYDIAAKFRGLTAVTNFDMNRYNVKATLESPLSGSAKR 393
DB 407 GNKOLYLGTFGTQBEAAAYDIAAKFRGLTAVTNFDMNRYNVKATLESPLSGSALARR 466
QY 394 LKEANRPVPSMMTNN--VSEENASGQWQNAVQHQQVDLS-----LLH 438
DB 467 -----NNNSIVVNTEDQTA--LNAVVEGGSNKVSTPERLLSPFAIPALP 510
QY 439 QHQRNYNGYNGNLS 455
DB 511 QVQKMFSGNM-GGNMS 526

RESULT 7
D86295
hypothetical protein T24D18.16 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
A;Accession: D86295
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H96827
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-308 <STO>
A;Cross-references: GB:AE005173; NID:g7715603; PIDN:AAF68121.1; GSPDB:GN00141
C;Genetics:
A;Gene: T20817.12
A;Map position: 1

Query Match 19.9%; Score 619.5; DB 2; Length 308;
Best Local Similarity 58.4%; Pred. No. 3.5e-36;
Matches 125; Conservative 31; Mismatches 41; Indels 17; Gaps 4;

QY 206 QRTSYRGVTRHRTGRYEAHLWNSCKREGQTRKGRQVYLGGYDKEEKAARAYDLAALK 265
DB 48 QRSSPYRGVTRHRTGRYEAHLWNSCKREGQTRKGRQVYLGGYDKEEKAARAYDLAALK 102
QY 266 YMGTTTTTTFPMSEYEVEEMKHTROYVAVSLRRKSSGFSRGASIVRGVTRHHQHGRW 325
DB 103 YWGRDTLLNPLPSYDVEDKEGQSKKEEYGSILRRKSSGFSRGVSKYRGVARHHNGRW 162

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D86295
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-332 <STO>
A;Cross-references: GB:AE005172; NID:g6587812; PIDN:AAF18503.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

Query Match 21.0%; Score 654.5; DB 2; Length 332;
Best Local Similarity 51.5%; Pred. No. 1.3e-38;
Matches 135; Conservative 42; Mismatches 66; Indels 19; Gaps 5;

QY 167 NSTSTCDN--NNDNNNVV--AQKTTIDDSVEATPKTIESFGORTSIYRGVTRHRTG 221
DB 8 SKTTSVENETGDDQATSVLKAKRRSQPRDAPP-----QRSSVHRGVTRHRTG 59
QY 222 RYEAHLWNSCKREGQTRKGRQVYLGGYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYE 281
DB 60 RYEAHLWNSCKREGQTRKGRQVYLGGYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYE 116
QY 282 KEVBEKMKHTROYVAVSLRRKSSGFSRGASIVRGVTRHHQHGRQWQARIGRVA 341
DB 117 EDIKEMESQKEEYTGSLRRKSSGFSRGASIVRGVTRHHQHGRQWQARIGRVA 176
QY 342 TFGTQBEAAAYDIAAKFRGLTAVTNFDMNRYNVKATLESPLSGSAKRLEKARPV 401
DB 177 TFGTQBEAAAYDIAAKFRGLTAVTNFDMNRYNVKATLESPLSGSAKRLEKARPV 233
QY 402 PSMMIMSNVSESENASGQWQ 423
DB 234 SPFIKPNHESDLSQSSSEDN 255

RESULT 8
H96827
protein P20B17.12 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
A;Accession: H96827
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H96827
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-308 <STO>
A;Cross-references: GB:AE005173; NID:g7715603; PIDN:AAF68121.1; GSPDB:GN00141
C;Genetics:
A;Gene: T20817.12
A;Map position: 1

Query Match 19.9%; Score 619.5; DB 2; Length 308;
Best Local Similarity 58.4%; Pred. No. 3.5e-36;
Matches 125; Conservative 31; Mismatches 41; Indels 17; Gaps 4;

QY 206 QRTSYRGVTRHRTGRYEAHLWNSCKREGQTRKGRQVYLGGYDKEEKAARAYDLAALK 265
DB 48 QRSSPYRGVTRHRTGRYEAHLWNSCKREGQTRKGRQVYLGGYDKEEKAARAYDLAALK 102
QY 266 YMGTTTTTTFPMSEYEVEEMKHTROYVAVSLRRKSSGFSRGASIVRGVTRHHQHGRW 325
DB 103 YWGRDTLLNPLPSYDVEDKEGQSKKEEYGSILRRKSSGFSRGVSKYRGVARHHNGRW 162

aintegumaenta-like protein - Arabidopsis thaliana
N/Alternate names: Protein T12E18.10
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C/Accession: T47591
R/Bloeker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat M.Mewes, submitted to the Protein Sequence Database, March 2000
A/Reference number: Z24469
A/Accession: T47591
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-205 <BLO>
A/Cross-references: EMBL:AL132971
A/Experimental source: cultivar Columbia; BAC clone T12E18
C/Genetics:
A/Map position: 3
A/Introns: 71/2; 99/1; 128/3; 170/2
A/Note: T12E18.10

Query Match	16.4%;	Score 511;	DB 2;	Length 205;
Best Local Similarity	55.4%;	Pred. No. 8.6e-29;		
Matches 103;	Conservative 23;	Mismatches 54;	Indels 2;	Gaps 2;

Qy 163 SLSMNSTSCDNNDS---NNNVAQKTIIDPSVEATPKKTIESFGQRTSIYRGVTRHRW 219
Db 15 SSSVSSSTTSSPIQSEAPRPKEAKRAKSSPSGDKSHNPTSPASTRRSSIYRGVTRHRW 74
Qy 220 TGYEHLWDNSCKRSGQTRKGQVYLGQYDKEEKAARAYDIALAALKYWGTTTTTNPMS 279
Db 75 TGFPEHLWDKSNWNIQNKKGQ---GAYDSEAAAHYDIALAALKYWGPDITLNPFAET 131
Qy 280 YKEVEVEEMKHTMQEYVASLRKSSGFSRGASTYRGVTRHHQHGRWQARIGRVAGNKDLY 339
Db 132 YTKLEEMQRTVKEEYLASLRQSSGFSRGVSKYRGVAVRHHNGRWEARIGRVFGNKLY 191
Qy 340 LGTFGT 345
Db 192 LGTYST 197

RESULT 11
T01574
indeterminate spikelet 1 - maize
C/Species: Zea mays (maize)
C/Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 29-Oct-1999
C/Accession: T01574
R/Chuck, G.; Mealey, R.B.; Hake, S.
Genes Dev 12, 1145-1154, 1998
A/Title: The control of maize spikelet meristem fate by the APETALA2-like gene indeterminate
A/Reference number: Z14353; MUID:98221107; PMID:9553044
A/Accession: T01574
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-433 <CHU>
A/Cross-references: EMBL:AF048900; NID:G2944039; PIDN:AAC05206.1; PID:G2944040
A/Experimental source: strain B73
C/Genetics:
A/Gene: idsl
A/Map position: 1

Query Match	15.2%;	Score 475.5;	DB 2;	Length 433;
Best Local Similarity	36.0%;	Pred. No. 7.7e-26;		
Matches 134;	Conservative 55;	Mismatches 116;	Indels 67;	Gaps 15;

Qy 118 SGSGCYGGDGGGSLGSMIKTWLRNPVDNVNONGNAAKGLS---LSMNSTSCDN 174
Db 16 SSSSVLSNADGG-----FRPGLLGSPVD--DDDCSGEMAPGASTGFMTQLPSPPTP 65
Qy 175 NDSNNNVA-----QKTIID-----SVEATPKKTIESFGQRTSIYRGVTRHRWTRGY 223
Db 66 PASEPEPVAAPVFWQVQPORADLGVQAQKVPAPAKVTRRGPSSRSQYRGVTFYRRTGRW 125
Qy 224 EHLWDNSCKRGQTRKGQVYLGQYDKEEKAARAYDIALAALKYWGTTTTTNPMSYEKE 283

Search completed: July 15, 2004, 10:03:41
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 09:52:29 ; Search time 13 Seconds
(without alignments)
2319.125 Million cell updates/sec

Title: US-09-980-364-2

Perfect score: 3119
Sequence: 1 MNNNWLGFLSPYEQNHRRK.....YHGGCGGEVAFPTFWNDN 579

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	461.5	14.8	432	1 AP2_ARATH	P47927 arabidopsis
2	150	4.8	248	1 AP2_ARATH	P47927 arabidopsis
3	145	4.6	248	1 PT16_LYCES	O04682 lycopersico
4	135.5	4.3	1146	1 AS10_YEAST	P48361 saccharomyc
5	132.5	4.2	344	1 RAV1_YEAST	Q92wm9 arabidopsis
6	132	4.2	243	1 ERF2_ARATH	O80338 arabidopsis
7	129.5	4.2	161	1 PT15_LYCES	O04681 lycopersico
8	126.5	4.1	1291	1 VAC4_HELPY	Q48258 heliocabacte
9	125	4.0	222	1 ERF1_ARATH	O80340 arabidopsis
10	124	4.0	268	1 ERF1_ARATH	O80337 arabidopsis
11	123	3.9	1341	1 YL78_YEAST	Q05854 saccharomyc
12	122.5	3.9	1253	1 D5PP_HUMAN	Q9nzv4 homo sapien
13	122	3.9	2339	1 RPCL_PLAFA	P27625 plasmodium
14	121.5	3.9	1296	1 VACL_HELPY	O48247 heliocabacte
15	121	3.9	961	1 GRP1_YEAST	Q2361 saccharomyc
16	120	3.8	964	1 CPA2_BACCI	P70873 bacillus ci
17	119.5	3.8	623	1 PNT1_DROME	P51022 drosophila
18	116.5	3.7	918	1 YMJ3_CABEL	P34487 caenorhabdi
19	116	3.7	1093	1 SWI4_YEAST	P25302 saccharomyc
20	115.5	3.7	880	1 BR4_DROME	Q42406 drosophila
21	114.5	3.7	1317	1 GPC_CABEL	P34288 caenorhabdi
22	113.5	3.6	1603	1 VIT3_CABEL	Q9n432 caenorhabdi
23	113.5	3.6	1603	1 VIT4_CABEL	P18947 caenorhabdi
24	113	3.6	300	1 ERF5_ARATH	O80341 arabidopsis
25	112.5	3.6	1290	1 VACA_HELPY	P55981 heliocabacte
26	112	3.6	719	1 NRP1_YEAST	P32770 saccharomyc
27	112	3.6	865	1 HTR3_ECOLI	P33129 escherichia
28	112	3.6	1310	1 VAC3_HELPY	O48253 heliocabacte
29	111.5	3.6	1603	1 VIT5_CABEL	P66125 caenorhabdi
30	111.5	3.6	1912	1 VIT1_CHICK	P87498 gallus gall
31	111	3.6	813	1 TTKA_DROME	P42282 drosophila
32	110.5	3.5	225	1 ERF3_ARATH	O80339 arabidopsis
33	110.5	3.5	416	1 FILA_HUMAN	P20930 homo sapien

34	110	3.5	405	1 YC8_YEAST	P25367 saccharomyc
35	110	3.5	578	1 PSF2_YEAST	P50109 saccharomyc
36	110	3.5	955	1 YCGV_ECOLI	P76017 escherichia
37	109	3.5	753	1 CKAA_BACUF	O32321 bacillus th
38	109	3.5	943	1 LBPA_NEIMB	Q06379 neisseria m
39	109	3.5	1087	1 POM1_SCHPO	Q09690 schizosacch
40	109	3.5	1541	1 IGA1_HAEIN	P42782 haemophilus
41	108.5	3.5	730	1 GLN3_YEAST	P18494 saccharomyc
42	108.5	3.5	1113	1 N116_YEAST	Q02630 saccharomyc
43	108.5	3.5	1211	1 BUN2_DROME	Q24523 drosophila
44	108	3.5	164	1 P202_LUPPO	P16146 lupinus pol
45	107	3.4	547	1 YCGV_ECO57	Q8X583 escherichia

ALIGNMENTS

RESULT 1

AP2_ARATH	STANDARD;	PRT;	432 AA.
AC	P47927		
DT	01-FEB-1996 (Rel. 33, Created)		
DT	01-FEB-1996 (Rel. 33, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Floral homeotic protein APETALA2.		
GN	AP2 OR AT4G36920 OR AP22.49 OR C7A10.440.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.		
OX	NCBI_TaxID:3702;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND MUTANTS AP2-1 AND AP2-5.		
RC	STRAIN=cv. Columbia; TISSUE=Flower;		
RX	MEDLINE=95003709; PubMed=7919989;		
RA	Jofuku K.D., den Boer B.G.W., van Montagu M., Okamoto J.K.;		
RT	"Control of Arabidopsis flower and seed development by the homeotic		
RT	gene APETALA2.";		
RL	Plant Cell 6:1211-1225(1994).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Columbia;		
RX	MEDLINE=20083488; PubMed=10617198;		
RA	Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,		
RA	Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,		
RA	Harris B., Ansgore W., Brandt P., Grivell L.A., Rieger M.,		
RA	Weichselgartner M., de Simone V., Obermaler B., Mache R., Mueller M.,		
RA	Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidheini T.,		
RA	Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,		
RA	Vos P., Hohnheisel J., Zimmermann W., Wedler H., Ridley P.,		
RA	Langanham S.-A., McCullagh B., Bilham L., Robben J.,		
RA	Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,		
RA	Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,		
RA	Weitzenecker T., Bothe G., Ramsperger U., Hilbert H., Braun M.,		
RA	Holzer E., Brandt A., Peters S., van Staveren M., Dirksen W.,		
RA	Koolman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,		
RA	Bernseiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,		
RA	De Keyser A., Buyschaert C., Giesen J., Villarroel R., De Clercq R.,		
RA	Van Montagu M., Rogers J., Cronin A., Quail M.A., Bray-Allen S.,		
RA	Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,		
RA	Pettitt A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,		
RA	Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,		
RA	Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,		
RA	Gabel C., Fuchs M., Fartmann B., Grandterth K., Dauner D., Herzl A.,		
RA	Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,		
RA	Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,		
RA	Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,		
RA	Chefor F., Cooke R., Berger C., Monfort A., Casacuberta E.,		
RA	Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,		
RA	Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,		
RA	Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,		
RA	Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker M.,		
RA	Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,		

FT	DNA_BIND	220	286	AP2/ERF 2.
FT	MUTAGEN	159	159	G->E: TEMPERATURE SENSITIVE; AP2-5.
FT	MUTAGEN	251	251	G->S: TEMPERATURE SENSITIVE; AP2-1.
FT	MUTAGEN	420	420	Q->E: TEMPERATURE SENSITIVE; AP2-5.
FT	SEQUENCE	432 AA;	47833 MW;	376A355291EFFB42 CRC64;
Query Match		14.8%;	Score 461.5;	DB 1; Length 432;
Best Local Similarity		34.3%;	Pred. No. 5.8e-24;	
Matches 114;	Conservative	59;	Mismatches 112;	Indels 47; Gaps 8;
QY	198	KXTIESFGCTGIYGVTRHRTGTGYEAHLWPNCKREGQTRKGRQVYLVGGYDKKEKAAR	257	
DB	119	KKSRCGRPSRSQYRGVTFYRTGWESHIND--C-----GQVILGGFDTAHAAR	168	
QY	258	AVDLAALKYGTWTTTTNTPMSYKEVEEMKMTQEOYVASLRRKSSGFSRGAISYRGVT	317	
DB	169	AYDRAAIKFRGYEADINFINIDYDDDLKQMTNLKEEFVHVLRROSTGPPRGSSKYRGVT	228	
QY	318	RHHQGRWQARIGRVAGNKDLVLGTFGTQEEAAEAYDIAAIKFRGLTAVTNFDMRYNVK	377	
DB	229	L-HKGRWEARWGQFLGKKYVLGLFDEVEAARAYDKAAIKCNKDAVTNFDPSIYDEE	287	
QY	378	AILESPLPIGSAAKRLKEANRPVPSMMIINNVSSENSASGWNAAVQHHQ-----	430	
DB	288	LNAES-----SGNPTTFQDENVLDLSLNSANSKHSQDMRLRMNQQQQDSLH	334	
QY	431	-----GVDLSSLHQQEYRNGYNGNGLSSARACFKQEDDQHP--LSNTQSLMTNI	483	
DB	335	SNEVLGLQGTGMLNHTPNSNQFPQSSNIGSGGFSLPFAAEN--HRFDGRASTNQVLTA	393	
QY	484	-----DHQSSVSDSDSVTCNVGVGVGYQ	507	
DB	394	RASSGFSPHHNNQINSTSTHQNLQTNGFQ	425	

RESULT 2

AP23 ARATH STANDARD; PRT; 248 AA.

AC P42736; O23105;

DT 01-NOV-1995 (Rel. 32, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE AP2 domain transcription factor RAP2.3 (Related to AP2 protein 3)

DE (Cadmium-induced protein AG30).

DE AT3G16770 OR MGL6.1 OR MGL6.24.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

NCBI_TaxID=3702;

[1]_RN

SEQUENCE FROM N.A.

RN STRAIN=cv. Columbia;

RN Choi S., Baek E., Jee S.;

RN Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.

RN [2]

SEQUENCE FROM N.A.

RN MEDLINE=97338147; PubMed=9192694;

RN MEDLINE=20277480; PubMed=10819329;

RN Okamura J.K., Caster B., Villarreal R., van Montagu M., Jofuku K.D.;

RN "The AP2 domain of APETALA2 defines a large new family of DNA binding

RN proteins in arabidopsis."

RN Proc. Natl. Acad. Sci. U.S.A. 94:7076-7081(1997).

RN [3]

SEQUENCE FROM N.A.

RN STRAIN=cv. Columbia;

RN MEDLINE=20277480; PubMed=10819329;

RN Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;

RN "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence

RN features of the regions of 4,504,864 bp covered by sixty P1 and TAC

RN clones.";

RN DNA Res. 7:131-135(2000).

RN [4]

SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;
RX MEDLINE=22954850; PubMed=14593172;
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Akrawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA Hayashizaki Y., Johnson-Hopson C., Hsuan W.W., Iida K., Karnes M.,
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamse R., Vaysberg M., Wallenberg E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinoraki K., Davis R.W., Theologis A., Becker J.R.,
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome."; 302:842-846 (2003).
RL Science
RN [5]
RN SEQUENCE FROM N.A.
RP Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Faldmann K.A.;
RA "Full-length cDNA from Arabidopsis thaliana";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- INDUCTION: By cadmium.
CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
CC -!- CAUTION: Ref.1 sequence is most probably a cloning artifact and
CC differs from that shown due to an insertion into the sequence.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z37504; CAA85734.1; ALT SEQ.
DR EMBL; AF003096; AAC49769.1; -
DR EMBL; AB022117; BAB02769.1; -
DR EMBL; AY059917; AAL24399.1; -
DR EMBL; AY035100; AAK59605.1; -
DR EMBL; AY142562; AAL13131.1; -
DR EMBL; AY087488; AAN65031.1; -
DR HSP; O80337; 2GCC.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR PRINTS; PR00367; ETHRSPELEMT.
DR PRODOM; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Cadmium.
FT DNA_BIND 76
FT DOMAIN 140 143
FT POLY-PRO.
SQ SEQUENCE 248 AA; 27758 MW; 317D50CC514782F1 CRC64;
Query Match 4.8%; Score 150; DB 1; Length 248;
Best Local Similarity 26.1%; Pred. No. 0.0027;
Matches 68; Conservative 38; Mismatches 105; Indels 50; Gaps 10;
QY 240 KGRVYLGVDYKEKAAAYDLAALKYGGTTTTFN-PMSEYEKEVEEMKHMTRQEVAS 298
DB 18 KGRKLTAEELWSELDAASADD-----FWGYSTSKLHPTNQVNVKEAVKKEQATPEGR 72
QY 299 LRRKSSGFSRGASTYRGVTRHQHGRWCARIGRVAGNKDLYLGTFTGQEEAAAYDAIAI 358
DB 73 RKRK-----NVYRGI-RKRPWKAAEIRDPRKGVRLVGLGTFTNTAEAAAYDVAAK 123
QY 359 KFRGLTAVTFN-DMNRVNVKALLESPLPTGSAAKLKEANRPVPSMMMSNNVSESENS 417
DB 124 QIRGDKAKLPD.LHHPPPPNYTPPPSSP-----RSTDQP-PAKKVCVQCSSELS 174
QY 418 ASGWQNAVQHQQGVDSLHLLHQHQRVYNGYNGNLSSESARACPKQEDDQHHLSNTQ 477

DB 175 QSPFVCEIGFNGDEPQ-----NLSYG-----FEPDYDLKQKQISSLE 212
QY 478 SLMT---NIDHQSSVSDSVT 495
DB 213 SFLELDGNTAQPSQLDESVS 233
RESULT 3
PT16_LYCSES
ID PT16_LYCSES STANDARD; PRT; 248 AA.
AC O04682;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pathogenesis-related genes transcriptional activator PT16.
GN PT16.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=97357308; PubMed=9214637;
RA Zhou J., Tang X., Martin G.B.;
RT "The Pto kinase conferring resistance to tomato bacterial speck
RT disease interacts with proteins that bind a cis-element of
RT pathogenesis-related genes";
RL EMBL J. 16:3207-3218 (1997).
CC -!- FUNCTION: Transcription factor that binds to the GCC-box
CC pathogenesis-related promoter element. Activates plants defense
CC genes.
CC -!- SUBUNIT: Interacts with the Pto kinase.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U89257; AAC49741.1; -
DR PIR; T07728; T07728.
DR HSP; O80337; 2GCC.
DR TRANSFAC; T04717; -
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR PRINTS; PR00367; ETHRSPELEMT.
DR PRODOM; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.
KW Transcription regulation; DNA-binding; Activator; Nuclear protein.
FT DNA_BIND 95 159
FT AP2/ERF.
SQ SEQUENCE 248 AA; 27913 MW; 115BD44DDEB1B992 CRC64;
Query Match 4.6%; Score 145; DB 1; Length 248;
Best Local Similarity 28.7%; Pred. No. 0.006;
Matches 50; Conservative 31; Mismatches 67; Indels 26; Gaps 7;
QY 144 NQVDNVNCE-NGNAAGLSLMSNSTSCNNNDNNNVVAQKTIIDSVEATPKTIE 202
DB 19 NKHFVSEHKEKNSLQVRVRIILTDATSSDDEGNTVRVRKHVTEINLMP--STK 76
QY 203 SFG--QRTSI-----YRGVTRHWRTGRYEAHLWDNSCKREGOTRGRQVYLG 249
DB 77 SIGDRKRERSVSDSVTRRKPRGVRQRPW-GRWAAEIRDPT-----RGRVWLGT 127
QY 250 DKEEKAARYDLAALKYWGTTTTTTFPMSEYEKEVEEMKHMTRQEVASLRKS 303
DB 128 DTPEEAAYVDKAAVKLKGPDATVTFPVST-TAEVTVTVTETETESVADGGDKS 180

```

RESULT 4
AS10_YEAST STANDARD; PRT; 1146 AA.
AC P48361;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ASK10 protein.
GN ASK10 OR YGR097W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / SEY6210;
RX MEDLINE=97060018; PubMed=8904339;
RA Page N., Sheraton J., Brown J.L., Stewart R.S., Bussey H.;
RT "Identification of ASK10 as a multicopy activator of Skn7p-dependent
RT transcription of a HIS3 reporter gene.";
RL Yeast 12:267-272 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Hernandez K., Weber N., Wipfli P., Schmidheini T.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Putative activator of SKN7.
CC -!- SIMILARITY: TO YEAST YIL105C AND YNL047C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U27209; AAA67368.1; -.
DR EMBL; U272882; CAA97100.1; -.
DR PIR; S64402; S64402.
DR GenOnline; 141409; -.
DR TRANSFAC; T03201; -.
DR SGD; S0003329; ASK10.
DR GO; GO:0006350; P:transcription; IGI.
DR InterPro; IPR001849; PH.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00233; PH; 1.
DR DOMAIN 22; 26 POLY-GLY.
DR DOMAIN 625; 628 POLY-SER.
DR DOMAIN 933; 938 POLY-ASN.
DR DOMAIN 958; 961 POLY-GLN.
DR DOMAIN 972; 975 POLY-SER.
DR CONFLICT 57; 57 I -> T (IN REF. 1).
DR CONFLICT 346; 346 T -> N (IN REF. 1).
DR CONFLICT 464; 464 P -> R (IN REF. 1).
DR CONFLICT 467; 467 V -> A (IN REF. 1).
DR CONFLICT 603; 603 A -> V (IN REF. 1).
DR CONFLICT 906; 906 L -> P (IN REF. 1).
SQ SEQUENCE 1146 AA; 126863 MW; 9E9EF22B0EE496EE CRC64;

Query Match
Best Local Similarity 4.3%; Score 135.5; DB 1; Length 1146;
Matches 114; Conservative 89; Mismatches 187; Indels 217; Gaps 31;

Qy 11 SPYQNHRRKDVYSTTTTVDVAGCYDPTAASDESSAIOISFPSPFGVVDATRDN 70
Db 507 SNYIHEFOSSDFYLNSSS-----TPNSTKSSAYSSVS-----IADTVANAN 548
Qy 71 NSHSRDWDINGACNNHNDQDQKLENFLGRITTTIYNINENVGDGSGCYGGGDDGG 130
Db 549 NAKA-----NNHHRQASD-----VHNSSTTT-----GGTAGAN 576

```

```

RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Militscher J., Miranda M., Nguyen M., Niemman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W., of the plant Arabidopsis
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RL thaliana."
RL Nature 408:816-820 (2000).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=22954850; PubMed=14593172;
RA Yamada K., Lim J., Dale J.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Southwick A.M., Wu H.C., Kim C.J., Lam B., Sakano H., Wu T., Yu G.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjani M., Hansen N.F.,
RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome."
RL Science 302:842-846 (2003).
CC -1- FUNCTION: Binds specifically to bipartite recognition sequences
CC composed of two unrelated motifs, 5'-CAACA-3' and 5'-CACCTG-3'.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: Expressed in all tissues examined: Roots,
CC rosette leaves, cauline leaves, inflorescence stems, flowers and
CC siliques. Highest expression in roots and rosette leaves. Very low
CC expression in flowers.
CC -1- DOMAIN: Contains two distinct DNA-binding domains. One is located
CC in the N-terminal region and binds to the 5'-CAACA-3' motif. The
CC second is located in the C-terminal region and binds to the
CC 5'-CACCTG-3' motif.
CC -1- SIMILARITY: Contains 1 AP2/ERF domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB013886; BAA34250.1; -
DR EMBL; AC011810; AAG09554.1; -
DR EMBL; AY063855; AAL36211.1; -
DR EMBL; AY091291; AAM14230.1; -
DR PIR; T51329; T51329.
DR HSSP; O80337; 2GCC.
DR InterPro; IPR003340; TF B3.
DR InterPro; IPR001471; TF ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR Pfam; PF02362; B3; 1.
DR PRINTS; PRO0367; ETRSPLENT.
DR PRODom; PD001423; TF ERF; 1.
KW DNA-binding; Nuclear protein.
FT DNA BIND 59 121 AP2/ERF.
FT DOMAIN 145 338 B3.
FT SEQUENCE 344 AA; 38597 MW; 7349B640B3505823 CRC64;
SQ
Query Match 4.2%; Score 132.5; DB 1; Length 344;
Best Local Similarity 21.7%; Pred. No. 0.064;
Matches 68; Conservative 48; Mismatches 125; Indels 73; Gaps 10;
Qy 177 DSNVNVVAGQKTTDDSVBATEPKKTIESFGQRTSIYGVTRHRTWGRYEAHLWDNSCKREG 236

```

```

Db 45 DSENGVEAESKLP-----SSKGVVPOP-NGRWCAQIYE----- 79
Qy 237 QTRGRQVYLGCDYDKEEKAAPAYDLAALKYWGTTTTTNF-PMSYEKEVEEMKMTQEQY 295
Db 80 ---KHORVWLGTFTNEDEEAARAYDVAVHFRFRDVTNFKDVQKMDDEVDVFLNSHSEI 136
Qy 236 VASLRKSSGFSRGASIVGYTRHHQHWCARIGRVAGNKDLYLGTFTGTEAAEAYDI 355
Db 137 VMLRKHTYNEBLEQS-----KRRNGN-----GNTRTLTSLGSLNDGVSTTGF 181
Qy 356 AAIKFRGLTAVTNFDMRNVKAI-----LESFSLPIGSAARLKEANRPVPSMM 405
Db 182 RSAEALFEKAVTPSDVGKLNELVLPKHAERKHPPLSSNVSVKGVLLNFEDVNGKWRFR 241
Qy 406 MISNVSSSENSASQWNAVOHH-----QGVDSLHLH-OHQERYNGYYNGG----- 452
Db 242 YSYWNSSOYVLTGWSRFVKEKNLRAGDVVYFSRSGDQDQLYIGMKSRSGLDAGR 301
Qy 453 -----NLSSESAR 460
Db 302 LRLFGVNISSPSSR 315
RESULT 6
ERP2_ARATH
ID ERP2_ARATH STANDARD; PRT; 243 AA.
AC O80338.
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Ethylene responsive element binding factor 2 (AtERF2).
GN ERF2 OR ERF-2 OR AT5G47220 OR MQL5.7.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20181733; PubMed=10715325;
RA Fujimoto S.Y., Ohta M., Usui A., Shinshi H., Ohme-Takagi M.;
RT "Arabidopsis ethylene responsive element binding factors act as
RT transcriptional activators or repressors of GCC box mediated gene
RL expression."
RL Plant Cell 12:393-404 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RL clones."
RL DNA Res. 7:31-63 (2000).
CC -1- FUNCTION: Transcription factor that binds to the GCC-box
CC pathogenesis-related promoter element.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Contains 1 AP2/ERF domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB008104; BAA32419.1; -
DR EMBL; AB018117; BAA97155.1; -
DR PIR; T51989; T51989.
DR HSSP; O80337; 2GCC.

```

DR TRANSFAC; T04634; --
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR PRINTS; PRO0367; ETRHSPLEMT.
 DR ProDom; PD001423; TF_ERF; 1.
 DR SMART; SM00380; AP2-1.
 KW Transcription regulation; DNA-binding; Nuclear protein.
 FT DOMAIN 23
 FT DNA_BIND 114 179 POLY-GLY.
 FT DOMAIN 192 206 AP2/ERF.
 FT DOMAIN 192 206 POLY-SER.
 SQ SEQUENCE 243 AA; 26797 MW; C9A4C9791249B5D4 CRC64;
 Query Match 4.2%; Score 132; DB 1; Length 243;
 Best Local Similarity 28.1%; Pred. No. 0.044; 53; Indels 22; Gaps 5;
 Matches 39; Conservative 25; Mismatches 53; Indels 22; Gaps 5;
 QY 192 SVEATPKKTI--ESFGQRTSIYRGVTRHRTGRYEHLWNSCKREGQTRKGRGVYLGYY 249
 DB 96 AMEEKPKKAIPVETAVAKHYRGVRQRPW-GKFAAEIRD-----PAKNGARVWLGTG 147
 QY 250 DKEEKAARAYDLAALKYWGTTTTTTFNPMSEYEKEVEEMKHTROEYVASLRKSSGFSRG 309
 DB 148 ETPAEDALAYDIAPFMGSRALLNFPURVNSGPPDVR-----ITSKRSSSSSSSS 200
 QY 310 ASIYRGVTRHHQGRQWAR 328
 DB 201 SS-----TSSSENGKLER 214
 RESULT 7
 ID PT15_LYCES STANDARD; PRT; 161 AA.
 AC O04681;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE Pathogenesis-related genes transcriptional activator PT15.
 GN PT15.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97357308; PubMed=9214637;
 RA Zhou J., Tang X., Martin G.B.;
 RT "The Pto kinase conferring resistance to tomato bacterial speck
 RT disease interacts with proteins that bind a cis-element of
 RT pathogenesis-related genes.";
 RL EMBO J. 16:3207-3218(1997).
 CC -!- FUNCTION: Transcription factor that binds to the GCC-box
 CC pathogenesis-related promoter element. Activates plants defense
 CC genes.
 CC -!- SUBUNIT: Interacts with the Pto kinase.
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U89256; AAC49740.1; --
 DR PIR; T07689; T07689.
 DR HSP; O80337; 2GCC.
 DR TRANSFAC; T04716; --
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 1.
 DR PRINTS; PRO0367; ETRHSPLEMT.

DR ProDom; PD001423; TF_ERF; 1.
 DR SMART; SM00380; AP2; 1.
 KW Transcription regulation; DNA-binding; Activator; Nuclear protein.
 FT DNA_BIND 56 121 AP2/ERF.
 SQ SEQUENCE 161 AA; 18051 MW; 67FAD2D4COB37FB0 CRC64;
 Query Match 4.2%; Score 129.5; DB 1; Length 161;
 Best Local Similarity 26.5%; Pred. No. 0.037; 66; Indels 41; Gaps 7;
 Matches 48; Conservative 26; Mismatches 66; Indels 41; Gaps 7;
 QY 137 MIKTLRNOPVNDVNDQENG-----NAAKGLSLMNSSTSCDNNDSNNNVVAQGTIDD 191
 DB 1 MYPTQSDIPLNENDSQEMVLYEVLENEALNIPVLPQR---NQLLPNNIL----- 49
 QY 192 SVEATPKKTI--ESFGQRTSIYRGVTRHRTGRYEHLWNSCKREGQTRKGRGVYLGYYDK 251
 DB 50 -----RPLQICIGKK---YRGVRRRPW-GKYAAEIRDSA-----RHGARVWLGTFT 91
 QY 252 EEKAAARAYDLAALKYWGTTTTTTFNPMSEYEKEVEEMK-----HMTROEYVASLRKSS 304
 DB 92 ABEAALAYDRAAFRMRGAKALLNFFSEIVNASVSDKLSLCSNYSYTTNNSSSLNEVSS 151
 QY 305 G 305
 DB 152 G 152
 RESULT 8
 ID VAC4_HELPY STANDARD; PRT; 1291 AA.
 AC Q48256;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Vacuolating cytotoxin precursor.
 GN VACA.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=185-44;
 RX MEDLINE=94335650; PubMed=8057855;
 RA Haas R., Schmitt W.;
 RT "Genetic analysis of the Helicobacter pylori vacuolating cytotoxin:
 RT structural similarities with the Iga protease type of exported
 RT protein.";
 RL Mol. Microbiol. 12:307-319(1994).
 CC -!- FUNCTION: Induces vacuolation of eukaryotic cells. Causes
 CC ulceration and gastric lesions.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z26883; CA881528.1; --
 DR PIR; S44983; S44983.
 DR InterPro; IPR006315; Autotransporter.
 DR InterPro; IPR005546; Autotransporter.
 DR InterPro; IPR003842; Vaca.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF02691; Vaca; 1.
 DR PRINTS; PRO1656; VACCYTOTOXIN.
 DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
 KW Toxin; Signal.
 FT SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 ? VACUOLATING CYTOTOXIN.

```
FT PROPEP ? 1291 POTENTIAL.
SQ SEQUENCE 1291 AA; 139635 MW; ECA56A61CAE36669 CRC64;

Query Match
  4.1%; Score 126.5; DB 1; Length 1291;
Best Local Similarity 18.0%; Pred. No. 0.97;
Matches 111; Conservative 73; Mismatches 191; Indels 241; Gaps 27;

QY 1 MNMNL-----GFLSPYEQNHRRKDVSTTTTVDVAGEYCYDPTAASDESSAIQTS 54
DB 254 MGNVMGRLOQVGYLAP-----SYSTINTSKVTGEVFNHLTVGDHNA----- 298
QY 55 FSPFGVGVDAFTRDNNSHRDWDI-----NGACNINHND 90
DB 299 ---QAGIIA-----SNKTHITGLDQWASAGLNIAPPBGYKDKPKDPSNTQNNAKD 350
QY 91 EQ-----DQPLENLFGRITTYNTNENNVGDSGG 121
DB 351 KQSSQNNSTQVNPNSAQTETQTOVIDP-----FAGGKNTVVNIRNTNADGII 406
QY 122 CYGGDG-----GGSLGLS-----MKTWLRNOPVD-----NVDNQNENNA 159
DB 407 RVGGFKASLTNAHLHIGKGGINLSQASGRSLLENLTGNTVDGPLRVNNGVGYAL 466
QY 160 KGLSLNNSSTCDNNDSN---NNVWAQKTIIDSVEATPKTIESFGORTSIYGVTRH 217
DB 467 AGSSANFEFKAGDTKNGTATFNNDISLGRFVNLKVA-----HTANFKGID-- 513
QY 218 RWTGRYEHLWDNSCKREGQTRKGRQVYLGQYDKEEKAARAYDLAALKYWGTTTTTFPM 277
DB 514 --TGN-----GGFN-----TLDFSGVTNKVNINK 535
QY 278 SEYEKEVEEMKHTQRYVNASLRKSGFSGRASIYRGVTRHHQHGWAQIRGV---AG 334
DB 536 LITASTNVAVKNFENELV-----KINGVSVE-----YTHFSEDIGSQSRINTVLETG 586
QY 335 NKDLYLGTGTOBEAAEYDIAIKFRLTAFTNFD-----MNRVYNKALLESPLP 386
DB 587 TRSYSG-----GVKPKGKLVINDFYAPNYFEDARNIKNVEITNKLA 631
QY 387 IGSAAKRLKEANRPVPMNMISNNVSENSA---SGWQNAAVQ---HHQGVDSLHLHQ 439
DB 632 FGQGS-----PWGTAKLFMNLITLQNAVMDYSQFSNLITQGFVANNQGT----- 677
QY 440 HQERYNGYNYGG---NLSSSARACPKQEDDQHHF-----LSNTQSLMTNIDHQS 488
DB 678 ---INLYRGQVATLVGNRAAMFFSNVDSATGFGYQPLMKINSADQLINKKEH--- 729
QY 489 VSDSVTVCGNVVGYG 504
DB 730 -----VLKAKIICYG 740

RESULT 9
ERF4 ARATH STANDARD; PRT; 222 AA.
ID AC O80340;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ethylene responsive element binding factor 4 (ATERF4).
GN ERF4 OR ERF-4 OR AT3G15210 OR K7L4.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20181733; PubMed=10715325;
RA Fujimoto S.Y., Ohta M., Usui A., Shirshi H., Ohme-Takagi M.;
RT "Arabidopsis ethylene responsive element binding factors act as
RT transcriptional activators or repressors of GCC box mediated gene
RT expression.";
```

```
RL RN
RP SEQUENCE FROM N.A.
RC STRAIN=cv Columbia;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger H., Perez-Alonso M., Obermaier B.,
RA Fartmann B., Valle G., Bloeker H., Maché R., Puigdomenech P.,
RA Delany M., Boutry M., Grivell L.A., Maché R., Puigdomenech P.,
RA De Simone V., Choisme N., Artiguenave F., Robert C., Brottier P.,
RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Quétier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Erle H., Jordan N., Bangert S.,
RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonacci B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,
RA Reichelt J., Schärfe M., Schoen O., Barges M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA Manthaupt G., Haase D., Schoof H., Rued S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Waits A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Waiti R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Miltscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Sasamoto S., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Kiyokawa S., Nakazaki N., Shimpō S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:820-822(2000).
CC -!- FUNCTION: Transcription factor that binds to the GCC-box
CC pathogenesis-related promoter element.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: Contains 1 AP2/ERF domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB008106; BAA32421.1; -.
DR EMBL; AF000413; BAB02150.1; -.
DR PIR; T52019; T52019.
DR HSP; O80337; 2GCC.
DR TRANSFAC; T04636; -.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 1.
DR PRINTS; PR00367; ETHRSEPELMNT.
DR ProDom; PD001423; TF_ERF; 1.
DR SMART; SM00380; AP2; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 22 86 AP2/ERF.
FT DOMAIN 132 137 POLY-GLY.
SQ SEQUENCE 222 AA; 23741 MW; CEF16ABA5ACFB073 CRC64;

Query Match
  4.0%; Score 125; DB 1; Length 222;
Best Local Similarity 33.6%; Pred. No. 0.11;
Matches 39; Conservative 12; Mismatches 39; Indels 26; Gaps 5;

QY 197 PKKTIESFQRTSI-YRGVTRHRTGRYEHLWDNSCKREGQTRKGROVYLGVDKEKA 255
DB 10 PATTNQTHNNAKEIRYGRKRPW-GRVAAEIRDPG-----KKTRVWLTFTAEHA 60
QY 256 ARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKHTQRYVNASLRKSGFSGRGA 311
DB 61 ARAYDTAARDPRGAKATNFPTF---LELSDQKVPT-----GPARSPS 100
```


DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Dentin sialoprotein precursor [Contains: Dentin phosphoprotein
 DE (Dentin phosphoprotein) (DPP); Dentin sialoprotein (DSP)].
 GN DSP.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=20168992; PubMed=10706475;
 RA Gu K., Chang S.R., Ritchie H.H., Clarkson B.H., Rutherford R.B.;
 RT "Molecular cloning of a human dentin sialoprotein gene.";
 RL Eur. J. Oral Sci. 108:35-42(2000).
 RN [2]
 RN SEQUENCE OF 463-1253 FROM N.A.
 RC TISSUE=Tooth;
 RX MEDLINE=99094526; PubMed=9879917;
 RA Gu K., Chang S.R., Slaven M.S., Clarkson B.H., Rutherford R.B.,
 RA Ritchie H.H.;
 RT "Human dentin phosphoprotein nucleotide and amino acid sequence.";
 RL Eur. J. Oral Sci. 106:1043-1047(1998).
 RN [3]
 RN DISEASE.
 RX MEDLINE=21096971; PubMed=11175779;
 RA Zhang X., Zhao J., Li C., Gao S., Qiu C., Liu P., Wu G., Qiang B.,
 RA Lo W.H.Y., Shen Y.;
 RT "DSP mutation in dentinogenesis imperfecta Shields type II.";
 RL Nat. Genet. 27:151-152(2001).
 RN [4]
 RN VARIANT=DFNA39/DG11 THR-17 AND PHE-18.
 RX MEDLINE=21096982; PubMed=11175790;
 RA Xiao S., Yu C., Chou X., Yuan W., Wang Y., Bu L., Fu G., Qian M.,
 RA Yang J., Shi Y., Hu L., Han B., Wang Z., Huang W., Liu J., Chen Z.,
 RA Zhao G., Kong X.;
 RT "Dentinogenesis imperfecta 1 with or without progressive hearing loss
 RT is associated with distinct mutations in DSPP.";
 RL Nat. Genet. 27:201-204(2001).
 CC -!- FUNCTION: DSP may be an important factor in dentinogenesis. DPP
 CC may bind high amount of calcium and facilitate initial
 CC mineralization of dentin matrix collagen as well as regulate the
 CC size and shape of the crystals.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed in teeth. DPP is synthesized by
 CC odontoblast and transiently expressed by pre-ameloblasts.
 CC -!- PTM: DSP is glycosylated.
 CC -!- DISEASE: Defects in DSP are the cause of dentinogenesis
 CC imperfecta 1 (DG11) (MIM:125490); also known as dentinogenesis
 CC imperfecta Shields type II (DG1-II). DG11 is an autosomal dominant
 CC disorder in which both the primary and the permanent teeth are
 CC affected. It occurs with an incidence of 1:8000 live births. The
 CC teeth are amber and opalescent, the pulp chamber being obliterated
 CC by abnormal dentin. The enamel, although unaffected, tends to
 CC fracture, which makes dentin undergo rapid attrition, leading to
 CC shortening of the teeth.
 CC -!- DISEASE: Defects in DSP are the cause of autosomal dominant
 CC deafness type 39 with dentinogenesis imperfecta 1 syndrome
 CC (DFNA39/DG11) (MIM:605594). Affected individuals present DG11
 CC associated with early onset progressive sensorineural high-
 CC frequency hearing loss.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF163151; AAF42472.1; -;
 CC EMBL; AF094508; AAD16120.1; -;

DR Genew; HNCN:3054; DSPP.
 DR MIM; 125485; -;
 DR MIM; 125490; -;
 DR MIM; 605594; -;
 KW Biomimetalization; Extracellular matrix; Signal; Glycoprotein;
 KW Phosphorylation; Sialic acid; Disease mutation; Deafness.
 FT SIGNAL 1 15
 FT CHAIN 16 1253
 FT CHAIN 16 462
 FT CHAIN 463 1253
 FT DOMAIN 439 1253
 FT SITE 488 490
 FT MOD_RES 259 259
 FT CARBOHYD 41 41
 FT CARBOHYD 43 49
 FT CARBOHYD 81 81
 FT CARBOHYD 130 130
 FT CARBOHYD 150 150
 FT CARBOHYD 190 190
 FT CARBOHYD 191 191
 FT CARBOHYD 209 209
 FT CARBOHYD 222 222
 FT CARBOHYD 275 275
 FT CARBOHYD 336 336
 FT CARBOHYD 387 387
 FT VARIANT 17 17
 FT /FTID=VAR_012280.
 FT V -> F (in DFNA39/DG11).
 FT /FTID=VAR_012281.
 FT MISSING (IN REF. 2).
 FT S -> SNGSDS (IN REF. 2).
 FT N -> D (IN REF. 2).
 FT S -> C (IN REF. 2).
 FT S -> G (IN REF. 2).
 FT MISSING (IN REF. 2).
 FT G -> S (IN REF. 2).
 FT N -> D (IN REF. 2).
 FT S -> G (IN REF. 2).
 FT N -> D (IN REF. 2).
 FT N -> N (IN REF. 2).
 FT G -> D (IN REF. 2).
 FT D -> E (IN REF. 2).
 FT E -> D (IN REF. 2).
 FT D -> E (IN REF. 2).
 FT D -> N (IN REF. 2).
 FT S -> R (IN REF. 2).
 FT CONFLICT 1132 1132
 FT CONFLICT 1132 1132
 FT SEQUENCE 1253 AA; 126426 MW; DAE240653904ED4A CRC64;
 Query Match 3.9%; Score 122.5; DB 1; Length 1253;
 Best Local Similarity 18.7%; Pred. No. 1.7; Indels 141; Gaps 21;
 Matches 102; Conservative 80; Mismatches 222;
 QY 43 AASDESSAIQTSPFPFGVYVDAFTDRNNSHSRDWDINGCACNNIHNDEQGPKEFLG 102
 Db 43 SYQDELNASGTIKES--GVLVHEGDRGRQENTQDG-----HKGEHGSKWAEEVG 90
 QY 103 RTTYYNTNENVDGSGCYGGCGGSGSLSMIKTWLRNQPVDVNDQNGNAAGL 162
 Db 91 KGFSTYTLAN--EGNIEG--WNGDTGKAETG-----HDGIGHKEENTANGI 136
 QY 163 ----SLSMNSSTSCDNNNDNNVVAQGTIDDSVEATPKKTIIESFGQRTSIYRGVTRR 218
 Db 137 QGVSIIDNAGATNRSNTNGTNDKNTQNGDVGADHEDVAVQEDGPQVAGSNNSTN- 195
 QY 219 WTGRVEHLWNSCKRBOGT-----RKGRVYLGVDYKKEKAAARVDAALKYWG 268
 Db 196 -----EDEIENSRCNEGTSBITPQINSKNGTK-----EAEVTPGTGDA-----G 238
 QY 269 TTTTTFNPFMEYKEVEEMKHTROYVYVASLRKKSFGSRGASIVRGVTRHHQGRWOAR 328
 Db 239 LNSDGSFSGNCADEDEDEGSCDDEBEAGKDKSSNNKQGE----- 281
 QY 329 IGRVAGNKDLYLGTGTGTEAEAAEADIAIKRPGTAVTNFDMNRYNKAILESPSLPIG 388

```
Db 282 -GDHGKEDDHDSSIGONSSKSYDP-----EGKEDPHNEVDGDKTSKSEENSACIPED 335
Qy 389 SAAKRL-----KEANRPVPMMTISNNVS--ESENSASGWAQVQHHQGVYDLSLLHQ 439
Db 336 NGSQRIEDTQKLNHRESKR-----VENRITKSETHAVG-----KSQDKGIEI----- 378
Qy 440 HQBRYNGYNGNLSESARACFKQEDDOHFL-----SNTQSLMTWIDHQS SVSDS 493
Db 379 -----KPPSSGNRNITKEVKGNGEGEDKGQHGMLGKGNVKTQGEVNIQGGQSEF- 432
Qy 494 VTCGNVVGYYGGVGGFAAPVNCDAASEFDYNARNHYFAQQQQTQQPGGDFPAAMTN 553
Db 433 ----GKNV-----GHSNTGSDNSDGYDSDYFD-----DKSQGDDPNSDE 470
Qy 554 NVGSN 558
Db 471 SNGND 475

RESULT 13
RFLC_PLAFA STANDARD; PRT; 2339 AA.
AC P27625;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA-directed RNA polymerase III largest subunit (EC 2.7.7.6).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92018020; PubMed=1656254;
RA Li W.B., Bzik D.J., Tanaka M., Gu H., Fox B.A., Inselburg J.;
RT "Characterization of the gene encoding the largest subunit of
RT Plasmodium falciparum RNA polymerase III.";
RL Mol. Biochem. Parasitol. 46:229-240(1991).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -!- SUBUNIT: RNA polymerase III consists of about 15 different
CC subunits. This subunit is the largest component of RNA polymerase
CC III.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are
CC found in eukaryotic nuclei: polymerase I for the ribosomal RNA
CC precursor, polymerase II for the mRNA precursor, and polymerase
CC III for 5S and tRNA genes.
CC -!- SIMILARITY: Belongs to the RNA polymerase beta' chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; M73770; AAA29729.1; --
DR InterPro; IPR000722; RNA_pol_A.
DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007066; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR InterPro; IPR006592; RNA_pol_N.
DR Pfam; PF04997; RNA_pol_Rpb1_1.
DR Pfam; PF00623; RNA_pol_Rpb1_2; 1.
DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
DR Pfam; PF05000; RNA_pol_Rpb1_4; 1.
DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.

DR SMART; SMO0663; RPOA.N; 1.
KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger; Nuclear protein.
FT ZN FING 88 101 C3H-TYPE (POTENTIAL).
SQ SEQUENCE 2339 AA; 272829 MW; EDD899363086FD48 CRC64;

Query Match 3.9%; Score 122; DB 1; Length 2339;
Best Local Similarity 15.1%; Pred. No. 4.3; Indels 300; Gaps 27;
Matches 105; Conservative 96; Mismatches 195;

Qy 69 DNNSRSDWDINGCACCNIHNDEQDGPKEFLGRTTITNTNENVVGSGSGCYGGD 128
Db 1103 NDSNMNSIHNNNSNMNIHNDSNRSIIHNDSNMNSIHNNDSNMNS----- 1150
Qy 129 GGSGLGLSMIKTWLENQPDVNDQENGNAAGLSLSMNSSTSCDNNDSNN----- 181
Db 1151 -----IHNNN-SNMNIHNDSNRSIIHNDSNMNSIHNNDSNMNSIHNNDSNMNS----- 1196
Qy 182 -----VVAQGTIDDSV-----EATPKKTIESFOQTSIYGVTR- 216
Db 1197 HNPYICNESLIIRNIMNELIYQNIQAQEDLFIPLEHDEFLVKNKIMESYTDQCNVEDIIRS 1256
Qy 217 -----HRWTGRY-----EHLWDSCK-----REGQTRK 240
Db 1257 LDLNKNVSYIHNDQGHLSLQWCAEEHITINNTNNTYVQEIMKLSKNKTKQSFK 1316
Qy 241 G-ROVYLGQYDKB-----KAARAY----- 259
Db 1317 GTIRDMH---EDSEQMKFKTKAKFFLEKKKGKHECDDIEYNTQYDNIQYNTSC 1373
Qy 260 -----DLAALKYWGTTTTTTFNPMSEYE-----KEYEEMKHMTRQ 293
Db 1374 NYIKSQNLNTHHQVNDLSPIK-----NNVILPPKEYSHFHFVNDYRNVEIKNLMDK 1428
Qy 294 EVV-----ASLRKSGSFGSGASIVRGVTRHHQH--GRWQARIGRVAGNKDLYL 340
Db 1429 KKFIFLNNEKNVQSKYRMSKNLKKIEIINNIYRNEKKLNKRWTKH---DNDDNY- 1483
Qy 341 GTFGTQEEAAEYDIAAI-----KFRGLTAVTFPMNRYNVKALLESPLFSGSAKRL 394
Db 1484 --WSSDDDSIIAKTIKIKKKEKRYHPKEKEKFNDRNYK----- 1522
Qy 395 KEANRPVPMMTISNNVSESENSASGWAQVQHHQGVYDLSLLHQHQRNGYNYN----- 450
Db 1523 -----MITDNNNNNNNNNNNNNNNNNNNN-----NNNNNNNNNNYNNLHDD 1563
Qy 451 -----GGNLSESAACFKQEDDOHFLSN-----TQSLMTWIDHQS VSD 491
Db 1564 VNNLGVNTYNTIYENDCNGIYEKETNNNELTNSMCDKNDPSEDFNNINENDLYD 1623
Qy 492 DSV--TVCGNVVGYGOGFAAPVNCDAASEFDY--NARNHYFAQQQQTQQSGGDF 547
Db 1624 NKYYRQIFKNVIGF-----VSVFEYVESYKQHYL-----F 1654
Qy 548 P-----AAMTNVGSNMYYHGGGGEVAPT 572
Db 1655 PYELIKWTSFLLEYLTEIPTNIFLTKLSKKEPT 1690

RESULT 14
VACL_HELPY
ID VACL_HELPY STANDARD; PRT; 1296 AA.
AC Q48247; Q53434;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Vacuolating cytotoxin precursor.
GN VACA.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
```

RP SEQUENCE FROM N.A.
 RC STRAIN=CCUG 17874 / NCTC 11638;
 RX MEDLINE=94222514; PubMed=8168917;
 RA Phadnis S.H., Ilver D.J., Janson L., Normark S., Westblom T.U.;
 RA "Pathological significance and molecular characterization of the
 RT vacuolating toxin gene of *Helicobacter pylori*";
 RL Infect. Immun. 62:1557-1565(1994).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RA MEDLINE=94216833; PubMed=8163943;
 RX Telford J.L., Ghara P., Dell'Orco M., Comanducci M., Burroni D.,
 RA Bugnoli M., Tecce M.P., Censini S., Covacci A., Xiang Z., Papini E.,
 RA Montecucco C., Parente L., Rappuoli R.;
 RA "Gene structure of the *Helicobacter pylori* cytotoxin and evidence of
 RT its role in gastric disease";
 RL J. Exp. Med. 179:1653-1658(1994).
 CC -!- FUNCTION: Induces vacuolation of eukaryotic cells. Causes
 CC ulceration and gastric lesions.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U07145; AAA18867.1; -
 DR EMBL; S72494; AAB30582.1; -
 DR InterPro; IPR006315; Autotransporter.
 DR InterPro; IPR005546; Autotransporter.
 DR InterPro; IPR003842; Vaca.
 DR Pfam; PF03797; Autotransporter; 1.
 DR Pfam; PF02691; Vaca; 1.
 DR PRINTS; PRO1656; VACCYTOTOXIN.
 DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
 KW Toxin; Signal.
 FT SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 ? VACUOLATING CYTOTOXIN.
 FT PROPEP 2 1296 POTENTIAL.
 SQ SEQUENCE 1296 AA; 139760 MW; 0D1F3F71AB411447 CRC64;
 Query Match 3.9%; Score 121.5; DB 1; Length 1296;
 Best Local Similarity 18.8%; Pred. No. 2.1;
 Matches 138; Conservative 83; Mismatches 224; Indels 289; Gaps 39;
 QY 1 MNNWNL-----GFSLSPEQNHKKVDVYSTTTTVDVAGEYCYDPTAASDESSAIGTS 54
 DB 254 MGNVWAGSLQVGVAYLAP-----SYSTINTSKVTGEVNFNHLTVGDKNA----- 298
 QY 55 FPSPEGVVDVDAFT-----RDNNSHRDWM-----DINGCANN 86
 DB 299 ---QAGIIANKKNTIGTLQWSAGLNIIAPPEGYKDKPNPTQSQAKNDKVESAKND 355
 QY 87 IHNDEQ-----DGPKEFLGRTTIIYNTNENVDGSGSG 121
 DB 356 KOESSQNSNTQVNPNSAQTVEQPTQVIDGP-----FAGGKDTVNNIRINNTNADGTI 411
 QY 122 CVGG-----GDGG-----GGSLGSLMIKTLWRNPQVD---NVDNCENG 156
 DB 412 RVGGFKASLTNAHLHGKGVNLNSQASGSL---IVENLTGNTIIVDGLRWNQVGG 468
 QY 157 NAAKGLSLSNMSSSTCDNNDSN--NNVAGKTIIDDSVEATPKKTTIESFGQRTSIYRGV 214
 DB 469 YALAGSSANFBFAKGTDTKNGTATFNNDISLGRFVNLKVD-----HTANFKGI 517
 QY 215 TRHRWTGRYEHLMDNSCKREGQTKRGVQLGVGYDKEERAAYDLAALKYWGTTTTTN 274
 DB 518 D-----TGN-----GGFN-----TLDFSGVTDKVN 537
 QY 275 FPMGEYEKEVEMKHTMTQEVYASLRKSSGFSRGASIIYRGVTRHHQHRWQARIGV-- 332

Db 538 INKLITASTNVAVKQNFNINELIV---KTNGISVGE-----YTHFSEDIGSOSRINTVRL 588
 QY 333 -AGNKDLYLTGTGTGTOBEAAEAYDIAAKFRG-----LTAVTNFDMRYNRYKAILE 381
 Db 589 ETGTRSLFSG-----GVKPKGKEKLVIDEFYSPWNYFDAR--NIKWEI 631
 QY 382 SPSLPISGSAKRLKEANRPVPMMSINNVSENSA---SQWQAAYQ-----HHQGVDL 434
 Db 632 TNKLAFPGQGS-----PWGTSKLMFNLTGLQAVMDYSQFSLNLTIOGFINNQT-- 682
 QY 435 SLLHQEORYNGVYVNGGNLSS-----ESARACFKOE-DDQHHF-----LSNTQSLMTNI 483
 Db 683 -----INLYRGKGVATLSVGNAAAMFNNDIDSATGYKPLKINSQAOLIKNT 732
 QY 484 DQSSVSDSVTVCGNVVGYG---GYOGFAAPVNCDAVAASEFD--YNARNHYFAQQCQ 538
 Db 733 EH-----VLLKAKIIGYGVNSTGTNGI--SNVNLEEQKFKERLALYNNNRMDTCVVRN 783
 QY 539 TQQSPG-----GDFPAAATNN-----VGS-----NMYVHG-----EG 565
 Db 784 TDDIRACGMAIGD--QSVVNPDPNYKILGKAWKNGISKANGSKISVYILGNSTPTEN 841
 QY 566 GGEVA--PTFTVMN 577
 Db 842 GGNNTNLPTNT--SN 855
 RESULT 15
 GPRI_YEAST STANDARD; PRT; 961 AA.
 ID GPRI_YEAST
 AC Q12361;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE G protein-coupled receptor GPRI.
 GN GPRI OR YDL035C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CC NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98190081; PubMed=9524122;
 RA Xue Y., Batille M., Hirsch J.P.;
 RT "GPRI encodes a putative G protein-coupled receptor that associates
 RT with the Gpa2p Galpha subunit and functions in a Ras-independent
 RT pathway";
 RL EMBO J. 17:1996-2007(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RX MEDLINE=97197972; PubMed=9046088;
 RA Saren A.M., Laamanen P., Lejarcegui J.B., Paulin L.;
 RT "The sequence of a 36.7 kb segment on the left arm of chromosome IV
 RT from *Saccharomyces cerevisiae* reveals 20 non-overlapping open reading
 RT frames (ORFs) including SIT4, FAD1, NAM1, RNA11, SIR2, NAT1, PRP9, ACT2
 RT and MFS1 and 11 new ORFs";
 RL Yeast 13:65-71(1997).
 CC -!- FUNCTION: Seems to associate with GPA2 and act as G protein-
 CC coupled receptor that senses glucose and controls filamentous
 CC growth. It acts upstream of adenylate cyclase and is required for
 CC glucose activation of cAMP synthesis in concert with a glucose
 CC phosphorylation-dependent mechanism.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- SIMILARITY: BELONGS TO FAMILY GPRI/GIT3 OF G-PROTEIN COUPLED
 CC RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC

```
CC or send an email to license@lsib-sib.ch).-----
CC EMBL; Z74083; CAA98593.1; -.
DR DR EMBL; Z71781; CAA96454.1; -.
DR PIR; S67568; S67568.
DR GarmOnline; 140277; -.
DR SGD; S0002193; GPR1.
DR GO; GO:0005886; C:plasma membrane, IDA.
DR GO; GO:0004930; F:g-protein coupled receptor activity; IMP.
DR GO; GO:0007186; P:g-protein coupled receptor protein signalin. . ; IMP.
DR GO; GO:0007124; P:pseudohyphal growth; IMP.
DR KW G-protein coupled receptor; Transmembrane.
FT TRANSMEM 57
FT TRANSMEM 79
FT TRANSMEM 92
FT TRANSMEM 114
FT TRANSMEM 134
FT TRANSMEM 156
FT TRANSMEM 179
FT TRANSMEM 198
FT TRANSMEM 251
FT TRANSMEM 273
FT TRANSMEM 620
FT TRANSMEM 642
FT DOMAIN 501
FT POLY-ASN 547
SQ SEQUENCE 961 AA; 110708 MW; 9889D857872A4209 CRC64;

Query Match      3.9%; Score 121; DB 1; Length 961;
Best Local Similarity 18.0%; Pred.No.1.5;
Matches 105; Conservative 83; Mismatches 201; Indels 194; Gaps 28;

Qy   9 SLSPYE---QNHRKDVYSTTTTVVDVACEYCVDPTAADESSAIQTSPSPFGVVVD 64
Db   :|:|:|:|:|:|:|:|:|:~VTFSVENEGLDTRKSMLGHTFSCNSLESPLAY-- 470
Qy   65 AFRDNHSHRDWPDIN----GCACNIHNDEQDPGLKENFLGRITTYINTNENVGDSG 119
Db   |:|:|:|:|:|:|:|:|:|:~NNKNNDNDNN 513
Qy   120 SGCVGGDDGGSGSLGMKTWLBNQPVDNVNQENGNAAGLSLVMSNSTSCDNNDSN 179
Db   -:|:|:|:|:|:|:|:|:|:~NNNNNNNNNNNN 541
Qy   180 NNVAQGKITDDSVVEAPPKTI-----ESFGQ-----RTSIYRGVTRHR 218
Db   NNRSNNIKVNVDNTNPADNIPILSNEAFTPSQQSQRVNNNADRCENSSFTNVQQH- 600
Qy   219 WTGRYEHLWDNSCKRGGOTRK-GQVLGYDYKEEKAARYDALALKYW----- 267
Db   :|:|:|:|:|:|:|:|:|:~YPLSYIGIWMLFPIADALQ 644
Qy   268 -----GTTTTTNF-----PMS-----EYEKEVEEMKHMTROYEVAS- 298
Db   645 YNHBIKGPTMVWTYYIDTCVPLESLCDVDVIIVLPFKEPWNYSWAKTESKYLI-EKYILKG 703
Qy   299 -----LRKSSGFSRGASIYGRTHRHQHGRWOARIGRVAGNKDLYLGTGTQEAAE 351
Db   704 ELGEKEILKFCHSNWGKRGWYR-----GWKKKRCWKYSTNPLKRIILWFVERPFKQ 755
Qy   352 AYDTIAATKFRGLTAVTFDM--NBRYNVKAILLESPLPGCSAAKLKEANRPVS-----M 404
Db   :|:|:|:|:|:|:|:|:|:~OGVDLSLLHQOERYNGYYYINGNISSESAR 460
Qy   756 LFEL-KLHFPSFYDCDDFEYEWNYSAKD--SDNKRITESDETNTSSDRSLPNSLBQ 812
Db   405 MMISNNVSESENSASGWAQAQQHH-----QGVDSLHLHQOERYNGYYYINGNISSESAR 460
Db   813 AMLNANIATAEEVPLFWR---IIHHIPMLGDIDLDELN---RLUKIYNNDHSFLPGLK 865
Qy   461 ACFAQK--EDDQHFFLSN-----TOSLMTNIDHQSVSD 492
Db   866 FALQNKSKDHQADVSTNSVMVKSGPPFNIVTN-DDENSIED 907
```

Search completed: July 15, 2004, 10:02:16
Job time : 16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 09:58:34 ; Search time 43 Seconds
(without alignments)
4248.489 Million cell updates/sec

Title: US-09-980-364-2

Perfect score: 3119

Sequence: 1 MNNNLGFLSLSYEQNHRRK.....YVHGEGGEVAFPTFTVNDN 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3119	100.0	579	10 Q8L3U3	Q8L3U3 brassica na
2	3041	97.5	579	10 Q8LSN2	Q8LSN2 brassica na
3	2484.5	79.7	584	10 Q8LSN1	Q8LSN1 arabidopsis
4	2475	79.4	581	10 Q9L3F42	Q9L3F42 arabidopsis
5	1139	36.5	655	10 Q7X6D4	Q7X6D4 oryza sativ
6	1131	36.3	692	10 Q8S1E1	Q8S1E1 oryza sativ
7	1089.5	34.9	597	10 Q8L3GQ3	Q8L3GQ3 oryza sativ
8	1040.5	33.4	516	10 Q9SYC2	Q9SYC2 arabidopsis
9	997.5	32.0	540	10 Q9LT37	Q9LT37 arabidopsis
10	997	32.0	566	10 Q9LXA3	Q9LXA3 arabidopsis
11	971.5	31.1	639	10 Q84Z02	Q84Z02 oryza sativ
12	963	30.9	492	10 Q8L3GQ4	Q8L3GQ4 oryza sativ
13	950.5	30.5	555	10 Q9F1E2	Q9F1E2 arabidopsis
14	930	29.8	485	10 Q41832	Q41832 zea mays (m
15	926	29.7	425	10 Q9CAH3	Q9CAH3 arabidopsis
16	922.5	29.6	437	10 Q9LSM4	Q9LSM4 arabidopsis

ALIGNMENTS

RESULT 1

Q8L3U3 PRELIMINARY, PRT; 579 AA.
ID Q8L3U3
AC Q8L3U3;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE AP2/EREP transcription factor BABY BOOM1.
GN BBM1.
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3708;
RN [1]_TaxID=3708;
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Topas;
RA Boutillier K., Offringa R., Sharma V.K., Kieft H., van Lammeren A.A.M.,
RA Quillet T., Zhang L., Hattori J., Liu C.-M., Miki B.L.A.,
RA Custers J.B.M., van Lookeren Campagne M.M.;
RT "Ectopic expression of the Brassica napus BABY BOOM gene triggers a
conversion from vegetative to embryonic growth.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF317904; AAM33800.1; -;
DR EMBL; AF317906; AAM33802.1; -;
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 2.
DR PRINTS; PR00367; ETRSPLENT.
DR ProDom; PD001423; TF_ERF; 2.
DR SMART; SM00380; AP2; 2.
SQ SEQUENCE 579 AA; 63962 MW; EA2APE4734500D72 CRC64;

Query Match 100.0%; Score 3119; DB 10; Length 579;

Best Local Similarity 100.0%; Pred. No. 8.6e-204; Indels 0; Gaps 0;
Matches 579; Conservative 0; Mismatches 0;

QY 1 MNNWLGFSLSPEQNHHRKDYVSTTTTVDVAGEYCYDPTAASDESSAIQTSPSPFG 60
 Db 1 MNNWLGFSLSPEQNHHRKDYVSTTTTVDVAGEYCYDPTAASDESSAIQTSPSPFG 60
 QY 61 VVVDFTDRNNSHSDWDINGCANNIHNDEODGPKLENFLGRITTTIYNTNENVDGSGS 120
 Db 61 VVVDFTDRNNSHSDWDINGCANNIHNDEODGPKLENFLGRITTTIYNTNENVDGSGS 120
 QY 121 GYGGGGGGGSLGLSMIKTWLRNPVDNDQENGAAGKLSLMSNSTSCDNNNDNN 180
 Db 121 GYGGGGGGGSLGLSMIKTWLRNPVDNDQENGAAGKLSLMSNSTSCDNNNDNN 180
 QY 181 NVVAQKTIIDSVATPKKTIIESFGORTSIYRGVTRHRTWGRYEHLWDNSCKREGOTRK 240
 Db 181 NVVAQKTIIDSVATPKKTIIESFGORTSIYRGVTRHRTWGRYEHLWDNSCKREGOTRK 240
 QY 241 GRQVYLGVDYDEEKAARAYDLAALKYWGTTTTTTPPKSEYEKEVEEMKHMTRQYVASLR 300
 Db 241 GRQVYLGVDYDEEKAARAYDLAALKYWGTTTTTTPPKSEYEKEVEEMKHMTRQYVASLR 300
 QY 301 RKSSGFSRGASIRYGVTRHHQGRWQARIIGRVAGNKDLYLGTFTGTOBEAAAYDIAAIKF 360
 Db 301 RKSSGFSRGASIRYGVTRHHQGRWQARIIGRVAGNKDLYLGTFTGTOBEAAAYDIAAIKF 360
 QY 361 RGLTAVTNFDMNRNVKAILSPISPIGSAARLKEANRPVPSMMTINNVSSENSASG 420
 Db 361 RGLTAVTNFDMNRNVKAILSPISPIGSAARLKEANRPVPSMMTINNVSSENSASG 420
 QY 421 WNAVAQHHQGVDSLHLLHQHORYNGYNGNLSSESARACFKQEDDOHFLSNTQSLM 480
 Db 421 WNAVAQHHQGVDSLHLLHQHORYNGYNGNLSSESARACFKQEDDOHFLSNTQSLM 480
 QY 481 TNIDHQSSVSDSVTCGVNMGVYGGYQGFAPVNCDAVAASEFDYNARNHYFFAQOQQTQ 540
 Db 481 TNIDHQSSVSDSVTCGVNMGVYGGYQGFAPVNCDAVAASEFDYNARNHYFFAQOQQTQ 540
 QY 541 QSPGGDFPAAMTNNVGSNNMYHGGGGGEVAPFTTWNNDN 579
 Db 541 QSPGGDFPAAMTNNVGSNNMYHGGGGGEVAPFTTWNNDN 579

RESULT 2
 QBLSN2
 ID QBLSN2 PRELIMINARY; PRT; 579 AA.
 AC QBLSN2;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE AP2/EREBP transcription factor BABY BOOM2.
 GN BMM2.
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3706;
 RN [1]_TaxID=3706;
 RP SEQUENCE FROM N.A.
 RC STRAIN=ev. Topas.
 RA Boutiller K., Offringa R., Sharma V.K., Kieft H., van Lammeren A.A.M.,
 RA Zhang L., Hattori J., Ouellet T., Liu C.-M., Miki B.L.A.,
 RA Custers J.B.M., van Lookeren Campagne M.M.;
 RT "Ectopic expression of the Brassica napus BABY BOOM gene triggers a
 RT conversion from vegetative to embryonic growth.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF317905; AAM33801.1;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2-domain; 2.
 DR PRINTS; PR00367; ETRSRLEMT.
 DR ProDom; PD001423; TF_ERF; 2.
 DR SMART; SM00380; AP2; 2.

SQ SEQUENCE 579 AA; 64018 MW; 6426D3B2B06D3BEA CRC64;
 Query Match 97.5%; Score 3041; DB 10; Length 579;
 Best Local Similarity 97.6%; Pred. No. 1.7e-198;
 Matches 565; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
 QY 1 MNNWLGFSLSPEQNHHRKDYVSTTTTVDVAGEYCYDPTAASDESSAIQTSPSPFG 60
 Db 1 MNNWLGFSLSPEQNHHRKDYVSTTTTVDVAGEYCYDPTAASDESSAIQTSPSPFG 60
 QY 61 VVVDFTDRNNSHSDWDINGCANNIHNDEODGPKLENFLGRITTTIYNTNENVDGSGS 120
 Db 61 VVVDFTDRNNSHSDWDINGCANNIHNDEODGPKLENFLGRITTTIYNTNENVDGSGS 120
 QY 121 GYGGGGGGGSLGLSMIKTWLRNPVDNDQENGAAGKLSLMSNSTSCDNNNDNN 180
 Db 121 GYGGGGGGGSLGLSMIKTWLRNPVDNDQENGAAGKLSLMSNSTSCDNNNDNN 180
 QY 181 NVVAQKTIIDSVATPKKTIIESFGORTSIYRGVTRHRTWGRYEHLWDNSCKREGOTRK 240
 Db 181 NVVAQKTIIDSVATPKKTIIESFGORTSIYRGVTRHRTWGRYEHLWDNSCKREGOTRK 240
 QY 241 GRQVYLGVDYDEEKAARAYDLAALKYWGTTTTTTPPKSEYEKEVEEMKHMTRQYVASLR 300
 Db 241 GRQVYLGVDYDEEKAARAYDLAALKYWGTTTTTTPPKSEYEKEVEEMKHMTRQYVASLR 300
 QY 301 RKSSGFSRGASIRYGVTRHHQGRWQARIIGRVAGNKDLYLGTFTGTOBEAAAYDIAAIKF 360
 Db 301 RKSSGFSRGASIRYGVTRHHQGRWQARIIGRVAGNKDLYLGTFTGTOBEAAAYDIAAIKF 360
 QY 361 RGLTAVTNFDMNRNVKAILSPISPIGSAARLKEANRPVPSMMTINNVSSENSASG 420
 Db 361 RGLTAVTNFDMNRNVKAILSPISPIGSAARLKEANRPVPSMMTINNVSSENSASG 420
 QY 421 WNAVAQHHQGVDSLHLLHQHORYNGYNGNLSSESARACFKQEDDOHFLSNTQSLM 480
 Db 421 WNAVAQHHQGVDSLHLLHQHORYNGYNGNLSSESARACFKQEDDOHFLSNTQSLM 480
 QY 481 TNIDHQSSVSDSVTCGVNMGVYGGYQGFAPVNCDAVAASEFDYNARNHYFFAQOQQTQ 540
 Db 481 TNIDHQSSVSDSVTCGVNMGVYGGYQGFAPVNCDAVAASEFDYNARNHYFFAQOQQTQ 540
 QY 541 QSPGGDFPAAMTNNVGSNNMYHGGGGGEVAPFTTWNNDN 579
 Db 541 HSPGGDFPAAMTNNVGSNNMYHGGGGGEVAPFTTWNNDN 579

RESULT 3
 QBLSN1
 ID QBLSN1 PRELIMINARY; PRT; 584 AA.
 AC QBLSN1;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE BABY BOOM.
 GN BMM.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN=C24;
 RA Boutiller K., Offringa R., Sharma V.K., Kieft H., van Lammeren A.A.M.,
 RA Ouellet T., Zhang L., Hattori J., Liu C.-M., Miki B.L.A.,
 RA Custers J.B.M., van Lookeren Campagne M.M.;
 RT "Ectopic expression of the Brassica napus BABY BOOM gene triggers a
 RT conversion from vegetative to embryonic growth.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF317907; AAM33803.1;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.

DR GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001471; TF_ERF.
DR Pfam: PF00847; AP2-domain; 2.
DR PRINTS: PR00367; ETRSPLEMT.
DR ProDom: PD001423; TF_ERF; 2.
DR SMART: SM00380; AP2; 2.
SQ SEQUENCE 584 AA; 64515 MW; 1EA3DCDFLC900PB8 CRC64;

Query Match 79.4%; Score 2484.5; DB 10; Length 584;
Best Local Similarity 81.3%; Pred. No. 1.2e-160;
Matches 474; Conservative 36; Mismatches 62; Indels 11; Gaps 6;

Qy 3 NNWLGFSLSPYEQNHHRKDVYSSTTTTVDVAGCYCDPTAADESSAIOQTSPPSPFGV 62
Db 5 NNWLGFSLSPHDQNHHRDVSSTTRTAVDVAGYCFDLAAPSDESSAVQTSFLSPFGVT 64

Qy 63 VDAFTRDNNSHSRDWDINGACNNIHNDQDQPKLENFLGRITTIYNTNENVDGSGSGC 122
Db 65 LEAFTRDNNSHSRDWDINGACNNIHNDQDQPKLENFLGRITTIYNTNENVDGSGC-C 123

Qy 123 YGGDGGGSGSLGSLMIKTWLRNOPVDNDQENGAAGLSLMSNSTSCDNNNDNNNV 182
Db 124 GGGDGGGSGSLGSLMIKTWLSNVANAHQDNGARGLSLMSNSTS-DSNNYNNDD 182

Qy 183 VAQKTIIDSVEATPKTIESFGQRTSIYRGVTRHRWTGRYEHLWDNSCKREGQTRKR 242
Db 183 VVQEKTIIVDVVETTPKKTIESFGQRTSIYRGVTRHRWTGRYEHLWDNSCKREGQTRKR 242

Qy 243 QVYLGVDKEEKAARAYDLAALKYWGTTTTNPFMSYEKEVEEMKHTROEYVASLRK 302
Db 243 QVYLGVDKEEKAARAYDLAALKYWGTTTTNPFMSYEKEVEEMKHTROEYVASLRK 302

Qy 303 SSFSGRSGASIRGVTRHHQGRQWARIQGVAGNKDLYLGTFTGQEEAAAYDLAAIKFRG 362
Db 303 SSFSGRSGASIRGVTRHHQGRQWARIQGVAGNKDLYLGTFTGQEEAAAYDLAAIKFRG 362

Qy 363 LTAVTFDMNRNRYNKAILESPLIGSAKRLKEANRPVPSMMISNNVSESENASGQ 422
Db 363 LSAVTFDMNRNRYNKAILESPLIGSAKRLKLVNVPVA-MMISNNVSESENANVSGWQ 421

Qy 423 NAAVQHHQGVDSLILHQQERYNGYNGNLSSESARACFKQEDDQHHFLSNTQSLMTN 482
Db 422 NTAFOHHQGMDSLILQQQERYVG-YINGNLSSTESTVCFKQEEQOHLRNSPSHMTN 480

Qy 483 IDHQSSVSDSVTCGNVVGYYGGYQGPAP-----VNCDAVAASEFDYNARNHYFA---Q 535
Db 481 VDHSSTSDSVTCGNVVGYYGGYQGPAPVGTSVNYDPPTAAEIAYNARNHYFAHQHQ 540

Qy 536 QOQTOQSPGDPFPAAMTNVGSNNMYHGGEGGEVAPTFTVWND 578
Db 541 QOQIQSPGDPFPVAISNNHSSNNMYFHGGEGGEGAPTFTSVWND 583

RESULT 4
Q9LF42 ID Q9LF42 PRELIMINARY; PRT; 581 AA.
AC Q9LF42; PRELIMINARY; PRT; 581 AA.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Ovule development protein aintegumenta-like protein.
GN T1086.90.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]

RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL391142; CAC01738.1; -.
DR PIR; T51580; T51580.
DR GO:0005634; C:nucleus; IEA.
DR GO:0003700; F:transcription factor activity; IEA.
DR GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001471; TF_ERF.
DR Pfam: PF00847; AP2-domain; 2.
DR PRINTS: PR00367; ETRSPLEMT.
DR ProDom: PD001423; TF_ERF; 2.
DR SMART: SM00380; AP2; 2.
SQ SEQUENCE 581 AA; 64170 MW; 292A8ACCC29BB79D CRC64;

Query Match 79.4%; Score 2475; DB 10; Length 581;
Best Local Similarity 81.3%; Pred. No. 5.3e-160;
Matches 474; Conservative 36; Mismatches 59; Indels 14; Gaps 7;

Qy 3 NNWLGFSLSPYEQNHHRKDVYSSTTTTVDVAGCYCDPTAADESSAIOQTSPPSPFGV 62
Db 5 NNWLGFSLSPHDQNHHRDVSSTTRTAVDVAGYCFDLAAPSDESSAVQTSFLSPFGVT 64

Qy 63 VDAFTRDNNSHSRDWDINGACNNIHNDQDQPKLENFLGRITTIYNTNENVDGSGSGC 122
Db 65 LEAFTRDNNSHSRDWDINGACNNIHNDQDQPKLENFLGRITTIYNTNENVDGSGC-C 123

Qy 123 YGGDGGGSGSLGSLMIKTWLRNOPVDNDQENGAAGLSLMSNSTSCDNNNDNNNV 182
Db 124 GGGDGGGSGSLGSLMIKTWLSNVANAHQDNGARGLSLMSNSTS-DSNNYNNDD 182

Qy 183 VAQKTIIDSVEATPKTIESFGQRTSIYRGVTRHRWTGRYEHLWDNSCKREGQTRKR 242
Db 183 VVQEKTIIVDVVETTPKKTIESFGQRTSIYRGVTRHRWTGRYEHLWDNSCKREGQTRKR 242

Qy 243 QVYLGVDKEEKAARAYDLAALKYWGTTTTNPFMSYEKEVEEMKHTROEYVASLRK 302
Db 243 Q---GGYDKEEKAARAYDLAALKYWGTTTTNPFMSYEKEVEEMKHTROEYVASLRK 299

Qy 303 SSFSGRSGASIRGVTRHHQGRQWARIQGVAGNKDLYLGTFTGQEEAAAYDLAAIKFRG 362
Db 300 SSFSGRSGASIRGVTRHHQGRQWARIQGVAGNKDLYLGTFTGQEEAAAYDLAAIKFRG 359

Qy 363 LTAVTFDMNRNRYNKAILESPLIGSAKRLKEANRPVPSMMISNNVSESENASGQ 422
Db 360 LSAVTFDMNRNRYNKAILESPLIGSAKRLKLVNVPVA-MMISNNVSESENANVSGWQ 418

Qy 423 NAAVQHHQGVDSLILHQQERYNGYNGNLSSESARACFKQEDDQHHFLSNTQSLMTN 482
Db 419 NTAFOHHQGMDSLILQQQERYVG-YINGNLSSTESTVCFKQEEQOHLRNSPSHMTN 477

Qy 483 IDHQSSVSDSVTCGNVVGYYGGYQGPAP-----VNCDAVAASEFDYNARNHYFA---Q 535
Db 478 VDHSSTSDSVTCGNVVGYYGGYQGPAPVGTSVNYDPPTAAEIAYNARNHYFAHQHQ 537

Qy 536 QOQTOQSPGDPFPAAMTNVGSNNMYHGGEGGEVAPTFTVWND 578
Db 538 QOQIQSPGDPFPVAISNNHSSNNMYFHGGEGGEGAPTFTSVWND 580

RESULT 5
Q7X6D4 ID Q7X6D4 PRELIMINARY; PRT; 655 AA.
AC Q7X6D4; PRELIMINARY; PRT; 655 AA.
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OSJNB0014K14.16 protein (OSJNB0116K07.8 protein).
GN OSJNB0014K14.16 OR OSJNB0116K07.8.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.

Qy	374	YNVKAILSPSLPIGSAKRLKEANRPVP-SMMISMNNVS-----ESENSASGW 421
Dd	444	YDVKSILDSAAALFVGTAAKRLKDFAAAAYVDVGRIASHLGGDCAIAAHYGHGHHHSAAAAM 503
Qy	422	QNAAVQ-----HGOGVDSLHLHOHERNGYYNYGCGNLSSSARACFKQEDDQHFLS 474
Dd	504	PRIAQAAPPPHAAG---LYHPYAQPLRGW-----C-KQEQD-HAVIA 543
Qy	475	NTQSMLTNIIDHOSSVDDSVTVCGNVGVGGYGQGAAPVNCDAVAASEFDYNARNHYFA 534
Dd	544	AASHLQ-DLHH-----LNLAGAAAA-----HDFFS 567
Qy	535	QQOQTQSQSGDGPAAATNNVGSN-MYYHGEGGG 567
Dd	568	QAWQQOHGLGSDNASLEHSTGNSVTYNGDGG 601
 RESULT 7 Q8LGQ3 PRELIMINARY; PRT; 597 AA.		
ID	Q8LGQ3	AC Q8LGQ3;
DT	01-OCT-2002	(TrEMBLrel. 22, Created)
DT	01-OCT-2002	(TrEMBLrel. 22, Last sequence update)
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)
DE	Ovule development	aintegumenta-like protein BNM3.
GN	BNM3	
OS	Oryza sativa (Rice).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
OC	Erbartoideae; Oryzaceae; Oryza.	
CX	NCBI_TaxId=4530;	
[1]	SEQUENCE FROM N.A.	
RN	Bi X.-Z.;	
RA	"Cloning and identification of two ovule development proteins,	
RT	aintegumenta-like protein in rice (Oryza sativa).";	
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AY062180; AAL47205.1; --	
DR	Gramene; Q8LGQ3; --	
DR	GO; GO:0005634; C:nucleus; IEA.	
DR	GO; GO:0003700; F:transcription factor activity; IEA.	
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.	
DR	InterPro; IPR001471; TF_ERF.	
DR	Pfam; PF00847; AP2-domain; 2.	
DR	PRINTS; PR00367; ETHRSPELENMT.	
DR	PRODOM; PD001423; TF_ERF; 2.	
DR	SMART; SM00380; AP2; 2.	
SQ	SEQUENCE 597 AA; 62198 MW; F856ERC99BADE25B CRC64;	
 Query Match 34.9%; Score 1089.5; DB 10; Length 597; Best Local Similarity 41.2%; Pred. No. 9e+66; Matches 264; Conservative 54; Mismatches 109; Indels 213; Gaps 20		
Qy	3	NNVLGRSLSPYQNHHRKDVYSTTTTVDVAGEVCYDPHTAASDESSAIOTSPSPFGVV 62
Dd	5	NNWLAFSLSPQCLP-----SQTNSTFIISAAA-----TTTTAGDSST----- 42
Qy	63	VDAFTRDNNSHSRDWDINGCACNTHNDQCPKLENFLGRITTIYNTNENVGDGSGGC 122
Dd	43	-----GDVCFNPQHPSTPAI----- 59
Qy	123	YGGDGGGSLGLSMIKTWLRNPQVDNVNQNGNAKGLSLSNWS----- 168
Dd	60	-----GNGIGLSMIKWLRSPQP-----QPAQALSLSMNMAAGTTTAOGGAMAL 105
Qy	169	-----STSC-----DNNDNSNNVVAOCKTIDDS-- 192
Dd	106	LACAGERGRTPASELSLTSIHGATTATWAGRKINEEGSGSAGAVVAVGSBSGSGAV 155
Qy	193	-----VEATPKKTIESFGORTSIYRGVTRHWRTGYEAHLMDNSCKREGOTRKGROYIG 247
Dd	166	VEAGAAAAAKRSVDFGORTSIYRGVTRHWRTGYEAHLMDNSCRREGOTRKGRO---G 222

248 GYDKEEKAARYDIALAALYKWTGTTTTTNPMSYEVEKEVEEMKHKHTROEYVASLRKSSGFS 307
 Db 223 GYDKEEKAARYDIALAALYKWTGTTTTTNPVNNYKELEEMKHKHTROEYVASLRKSSGFS 282
 Qy 308 RCASIYRGVTRHHOHGRWOARIGRVAGNKDYLGTGTQEEAAEYDIAAIKPRGLTAVT 367
 Db 283 RCASIYRGVTRHHOHGRWOARIGRVAGNKDYLGTGTQEEAAEYDIAAIKPRGLNAV 342
 Qy 368 NFDNRYNNKAILSPSPILPTIGSAKRLKEANRPVP-SMMISNNVS-----ESE 415
 Db 343 NFDMSRYDVKSILDSAAALPVGTAAKRLKDAEAAAYDVGRIASHLGGDGGAYAAHYGHHH 402
 Qy 416 NIASGWQNAAVQ-----HHQGVDLILLHOHORVNGYNYNGNLSSESARACFKQED 468
 Db 403 SAAAWPPIAFCAAAAPPPHAG---LYHPYAQPLRGW-----C-KQED 443
 Qy 469 QHHEFNTQSMTNIDHOSSVDSVTCGVNVGYGCGYGAFAVNCDAVAASEFDYNAR 528
 Db 444 -HAVTAAAHSLQ-DLHH-----LNLGAAAAA- 467
 Qy 529 NHYVFAQQOQOQSQSGGDFPAAATNNVGSN-MYVHGEGGG 567
 Db 468 -HDFSQAMQOQHGLGSDNASLHSTGSGSVVYNGDNGG 506
 RESULT 8
 Q9SYC2 PRELIMINARY; PRT; 516 AA.
 AC Q9SYC2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE F1M15.6 protein.
 GS F1M15.6.
 CS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosid1; brassicales; Brassicaceae; Arabidopsis.
 NC CNE1_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
 RA Altafi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
 RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
 RA Luros S., Schwartz J., Shinn P., Toriumi M., Vystotskaia V.S.,
 RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.,
 RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AC006085; AAC30633.1; -
 DR FBL; F96549; F96549.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001471; TF_ERF.
 DR Pfam; PF00847; AP2_domain; 2.
 DR PRINTS; PR00367; ETHRSPLEMT.
 DR ProDom; PD001423; TF_ERF; 2.
 DR SMART; SM00380; AP2; 2.
 SQ SEQUENCE 516 AA; 56335 MW; 80996A95851B8BEA CRC64;
 Query Match 33.4%; Score 1040.5; DB 10; Length 516;
 Best Local Similarity 44.6%; Pred. No. 1.6e-62;
 Matches 254; Conservative 70; Mismatches 114; Indels 131; Gaps 229
 Qy 86 NIHNDEQDC---PKLENFLGRITTT-IYNTNEN-----VGDGSGCGYGGDGGGSLGL 116
 Db 3 NPHGGGEGGEPKVAADFLGVSKSGDHHDTNHLVPYNDIHQTNASDYFQTN---SL-L 57
 Qy 136 SMIKTWLRNPVDVNDQENGNAAGKLSLNNSS-----TSCDNNDNS 179
 Db 58 PTVVTCASNAP--NNVELQESAHTNQSLTSLMGSTGAAAAEAVATVKASPATSDNDSST 229
 Qy 180 NNVAQCKTIDDSVEATPKKTIESFGQRTSYIRGVTHRTTGRYEALWNDSCKRGQTR 239

```

Db 117 N---TSGAI---VATPRILETPGQTSIYRGVTRHRWTRGVEAHLWDSNCRREGQSR 170
QY 240 KGRQVYLGDDYDKEEKAARAYDLAALKYWGTTTTNPFPMSEYEKEVEEMKHMTRQEYVASL 299
Db 171 KGRQ---GGYDKEEKAARAYDLAALKYWGPTTTFITNYEKEVEEMKHMTRQEFVASI 227
QY 300 RKSSGFRGASIIYGVTEHGHGHWQARI GRVAGNKDLYLGTCTOEAREAEAYDIAIK 359
Db 228 RKSSGFRGASMYGVTRHHGHGHWQARI GRVAGNKDLYLGTCTOEAREAEAYDIAIK 287
QY 360 FRGLTAVTNFDMNRNVNKAILESPSLPI-GSAAKRLKEA-----NRVPVSMMSNNVSE 413
Db 288 FRGLNAVTFEINRYDVKAILESNLPIGGGAKEAQLKEAQLKEAQLKEAQLKEAQLKEAQLKEA 347
QY 414 SENSASGQNA---VOHQGV-----DLS----- 435
Db 348 -YGAASGSSVASSSRRLQLPYPLSIQPFELHHLHQLPPLLQNNNDISQYHDSFSYIQT 406
QY 436 LLHOHOERYNGYNGNLSSESARACFKQBDODHHFLSNLQSLMTNIDHQSVDSDSVT 495
Db 407 QLHLHQQTNYLQSSHTSQL-----YNAVLSQNPGLL-----HGFVSDNN-- 448
QY 496 VGNVVGYGQGFAPVNCDAAYASEFDYNARNHYFAOQQOQTOQSPGGDFPAAMTNNV 555
Db 449 ---NTSGFLNNGIGIGSSSTVGSABEEFPVAVKDY-----DMPPS----- 487
QY 556 GSNMYVHGGEGGAVP-----TFTVND 578
Db 488 GGATCYGWNSSGESAGSNPGGVFTWNE 516

RESULT 9
Q9LX3
ID Q9LX3 PRELIMINARY; PRT; 540 AA.
AC Q9LX3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similarity to unknown protein.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RA Nakamura Y.;
RX MEDLINE=20277480; PubMed=10819329;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:131-135(2000).
DR EMBL; AB025629; BAB02492.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 2.
DR PRINTS; PR00367; ETRHSPLEMTN.
DR ProDom; PD001423; TF_ERF; 2.
DR SMART; SM00380; AP2; 2.
SQ SEQUENCE 540 AA; 59066 MW; 0A013B0735176DDD CRC64;

```

Query Match 32.0%; Score 997.5; DB 10; Length 540;
 Best Local Similarity 40.6%; Pred. No. 1.4e-59;
 Matches 239; Conservative 80; Mismatches 142; Indels 127; Gaps 18;

```

QY 69 DNNHSRDWDINGCACNNIHNDEQDQ-----PKLENFLGRITTIYNTNENVGDSGSGCYG 124
Db 2 DNPFTQENWM-----INPHGGGDEGEVVPKADFLGVSKPDENQSHLVAYNDSDYF 56
QY 125 GGDGGGSLGLSMIKTLRNQPV-----DNVDN-----QENGNAAGLSLMSNSTSC 172
Db 57 HTN-----SLMPSVQSDNVVAAACDSNTFNSSYHQLQESAHNLSLQSLTSMGTT--- 105
QY 173 DNNNSNNNV---AQGKTIDDSVE-----ATPKKTIESFQRTSIYRGVTRHRWT 220
Db 106 ----AGNNVDPKASPSSETTGDNASGGALAVVETATPRALDTFCQRTSIYRGVTRHRWT 160
QY 221 GRYEHLWDSNCKRGQTRKGRQVYLGVDYDKEEKAARAYDLAALKYWGTTTTNPFPMSEY 280
Db 161 GRYEHLWDSNCKRGQTRKGRQVYLGVDYDKEEKAARAYDLAALKYWGTTTTNPFPMSEY 217
QY 281 EKEVEEMKMTQRYEYASLRKSSGFRGASIIYGVTRHRHGHGHWQARI GRVAGNKDLYL 340
Db 218 EKEVEEMKMTQRYEYASLRKSSGFRGASIIYGVTRHRHGHGHWQARI GRVAGNKDLYL 277
QY 341 GTFGTEBAEAAYDIAIKERGLTAVTNFDMNRNVNKAILESPSLPI-GSAAKRLKEA-- 397
Db 278 GTFTEBAEAAYDIAIKERGLTAVTNFDMNRNVNKAILESPSLPI-GSAAKRLKEAQA 337
QY 398 -----NRVPVSMMSNNVSE 413
Db 338 LESSRKRKAEMIALGSSFOYGGSGSTGSSRLQLQPYPLSIQPLEPFLSLQNNDIS 397
QY 414 SENSASGQNA---VOHQGV-----DLS----- 435
Db 398 HYNNNAHDSSTFNHSHYIQ-TQLHLHQQTNYLQSSHTSQL-----YNAVLSQNPGLL-----HGFVSDNN-- 456
QY 474 SNTQSLMTNIDHQ---SSVSDSDSVTCVNVGVGGYQGFAPV-NCDAAYASEFDYNARNH 530
Db 457 VST-SIVDNNNNNGSSGSSYNATAFLGN---HGIGIGSSSTVGSSTEEPTVKTDYD---- 508
QY 531 YVFAOQQOQTOQSPGGDFPAAMTNNVGSNNYHGGEGGAVPFTVND 578
Db 509 -----MPSSDGTGYSCTWSVSGSNPGG-----VFTWNE 540

RESULT 10
Q9LX3
ID Q9LX3 PRELIMINARY; PRT; 566 AA.
AC Q9LX3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Orule development protein-like.
GN F12B17_140.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL335395; CAB89392.1; -.
DR F12B17_140.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_ERF.
DR Pfam; PF00847; AP2-domain; 2.
DR PRINTS; PR00367; ETRHSPLEMTN.
DR ProDom; PD001423; TF_ERF; 2.
SQ SEQUENCE 566 AA; 59066 MW; 0A013B0735176DDD CRC64;

```

```
DR SMART: SMO0380; AP2; 2.
SQ SEQUENCE 566 AA; 63257 MW; 470627FB3002744A CRC64;

Query Match 32.0%; Score 997; DB 10; Length 566;
Best Local Similarity 40.4%; Pred. No. 1.6e-59;
Matches 247; Conservative 58; Mismatches 127; Indels 180; Gaps 18;

QY 43 AASDESSAIOFSFSPGCVVDAETRDNNSHSRDWDINGCACNNIHNDQDGPGLNPLG 102
DB 52 AASVADSTILTFVDP-----OSHSQNHIPKLEDFLG 84

QY 103 RT-----TTIY-----NTNEN 113
DB 85 DSSSIVYSDNSQTDQDSSLTQIYDPRHHNQTFYSDHDDFKTMAGFQSAFSTNSGSE 144

QY 114 VGDGSGG-----CYGGDGGGSLGSLMIKTLWLNQPVNDVNDQ 153
DB 145 VDDSGASGTRHLAGDYLGHVYESSGPGLGFHGGSTGALSLGVNNNN--TNHRDN-DNH 201

QY 154 ENGNAAGLSLMSNSTSCDNNNDNNNVVAQK---TIDDSVEATPKKTIESFGQRTS 209
DB 202 YRGN-----NNGERINNNNDNEKTDSEKAVAVAVETSDCSNKKIADTFGQRTS 252

QY 210 IYRGVTRHRTGRYEHLWNSCKREGQTRKGRQVYLGVDKBEKAARAYDLAALKYWG 269
DB 253 IYRGVTRHRTGRYEHLWNSCRREGQARKGRQ---GGYDKEDKAARAYDLAALKYWNA 309

QY 270 TTTNFPMSYEKEVEEMKMTROYVASLRKSSGFSRGASIVRGVTRHHQGRWOARI 329
DB 310 TATTNFIITNYSKEVEEMKMTKOEFTASLRKSSGFSRGASIVRGVTRHHQGRWOARI 369

QY 330 GRVAGNKDLYLGTGTQEEAAEAYDIAAKFRGLTAVTNFDMNRYNKAILESPLSIG 389
DB 370 GRVAGNKDLYLGTGTATBEAAEAYDIAAKFRGINAVTNFEMNRYDVEALMK-S-ALPIGG 428

QY 390 AAKELKEANRPVPSMMNISNNVSESENASGQWNAAVQHOGVDLSLHGHQRYNGYYY 449
DB 429 AAKELK-----LSLEAAASSEOKFILGHQ-----LHFQOOQ-----461

QY 450 NGGNLSSESARACPKQEDDQHHFLSNTQSLMTNIDHQSVDSDSVTCGVNVVGYGQGF 509
DB 462 -----QQOQLQSSPNHSSINFALCPNSAVSQSIIFCG-----496

QY 510 AAPVNCDAAYASEFDYNARHHYFAQQOCTQSPGGFPA-AMTNVGSNNYHVGEGG-G 567
DB 497 ---IPFEAALYHHHQOQOQ-----QQOQOQOQFFQHFPAANAASDSTGSNNNSNVQTMG 549

QY 568 EVAPT---FTW 576
DB 550 LMAENPAEFFLM 561

RESULT 11
Q84202 ID Q84202 PRELIMINARY; PRT; 639 AA.
AC Q84202;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 25, Last annotation update)
DE Putative ovule development protein antitegumenta (ANT).
GN P0474G09.16.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC
clone: P0474G09.16";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF005309; BAC56815.1; -.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001471; TF_ERF.
DR Pfam: PF00847; AP2-domain; 2.
DR PRINTS: PR00367; ETRSPLEWMT.
DR ProDom: PD001423; TF_ERF; 2.
DR SMART: SMO0380; AP2; 2.
SQ SEQUENCE 639 AA; 67622 MW; EE4BC5123C56E68F CRC64;

Query Match 31.1%; Score 971.5; DB 10; Length 639;
Best Local Similarity 36.3%; Pred. No. 1e-57;
Matches 252; Conservative 79; Mismatches 174; Indels 189; Gaps 22;

QY 3 NNWLGSLSPY-----BQNHHRKD-----VYSSTTTTVDVAVAG 35
DB 16 SGNLGSLSFPHMASSTMDSEHHVHHHQQOQQOQQOQQHQQOQGLFPFVPTAAAAAY 75

QY 36 EYCYDPTASDESSAIOFSFSPG---VVVDAFTRDNNSHSRDWDINGCACNNIHNDQ 92
DB 76 GLAGDVVAATNGYYSQLASMPLKSDGSLCIMEALRRTDQDH-----116

QY 93 DQPKLENPLG-----RTTIYNTNENVDGSGGCGYGGDGG-----GSLGLSMI- 138
DB 117 HGFPLEDFLGAAQAPAMALSLDNTSGFYGGGCAAAAGHGHQGYLOACDLYGGPAAPSLVT 176

QY 139 -----KWL-ENQPDVNDVQENGAAK-----GLSLMN 167
DB 177 ADEEAAAAAAMAAVAAARGAATAATGAADANAENVLPSATAAQHLHPPLALSMSSG 236

QY 168 SSTSDNNNDNNNVVA-----QGKTIDDSVEATPKKTIESFGQRTSIYR 212
DB 237 SLSSCITAGEYGAATAAADGGRKGGAGGQKQPVHH-----RKSIDTFGQRTSQYR 290

QY 213 GVTTRHRTGRYEHLWNSCKREGQTRKGRQVYLGVDKBEKAARAYDLAALKYWGTTT 272
DB 291 GVTTRHRTGRYEHLWNSCKREGQTRKGRQ---GGYDMEEKAAARAYDLAALKYGPSTH 347

QY 273 TNFPMSYEKEVEEMKMTROYVASLRKSSGFSRGASIVRGVTRHHQGRWOARI 332
DB 348 INFPLEDQEELEEMKMTROYVAHLARKSSGFSRGASIVRGVTRHHQGRWOARI 407

QY 333 AGNKDLYLGTGTQEEAAEAYDIAAKFRGLTAVTNFDMNRYNKAILESPLSIGSAK 392
DB 408 SGNKDLYLGTGTQEEAAEAYDIAAKFRGLNAVTFDITRYDVKIWAASNTLLPADLAR 467

QY 393 RLKEANRPVPSMMNISNNVSESENASGQWNAAVQHOGVDLSL-----HGHQ 442
DB 468 R-----NAATTSKDDHSAAG-AGAIVSVHSAADIADVADTLWKATAPRQOQ 514

QY 443 RYNGYYINGN-----LSSESARACPKQEDDQHHFLSNTQSLMTNIDHQSVDSD 492
DB 515 HEDDVVLSGADQAFSVLHDLVAVDAAAHHQOQQOQQHMSMSAASSLVTSLNSREGSPD 574

QY 493 SVTVCGNVVGYGQGFPAAPVNCDAAYASEFDYNARHHYFAQQOCTQO-----SPG 547
DB 575 R-----GGLSMLFAKP--SPAFAAS-----AQOQASTKLMAAPLPLGWS 612

QY 548 ---PAAMTNVGSNNYHVGEGGGEVAPTFTVMD 578
DB 613 VSSPPASAPPPAVSIH-----MPLFAAWTD 638

RESULT 12
Q8LQ04 ID Q8LQ04 PRELIMINARY; PRT; 492 AA.
AC Q8LQ04;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Aintegumenta-like protein.
OS Oryza sativa (Rice).
```

1

QY	497	CGN-----VVGVGVGQGFAPFVNCDAVAASEFDYNARNHY	532
DB	436	ASNPYAAAVVSGTAGYEGTGNGTWVTITSSNTGTAPHY	478
RESULT 15			
Q3CAH3	Q3CAH3	PRELIMINARY;	PRT; 425 AA.
ID	Q9CAH3		
AC	Q9CAH3;		
DT	01-JUN-2001	(TREMELrel. 17, Created)	
DT	01-JUN-2001	(TREMELrel. 17, Last sequence update)	
DT	01-JUN-2003	(TREMELrel. 24, Last annotation update)	
DE	Putative AP2 domain transcription factor.		
GN	F2BP22.24		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;		
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.		
OX	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Columbia;		
RX	MEDLINE=21016719; PubMed=11130712;		
RA	Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,		
RA	White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,		
RA	Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,		
RA	Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,		
RA	Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,		
RA	Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,		
RA	Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,		
RA	Kim C.J., Koo H.B., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,		
RA	Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,		
RA	Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,		
RA	Militischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,		
RA	Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,		
RA	Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,		
RA	Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,		
RA	Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,		
RA	Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;		
RT	"Sequence and analysis of chromosome 1 of the plant Arabidopsis		
RT	thaliana."		
RL	Nature 408:816-820(2000).		
DR	EMBL; AC010926; AAGS1860.1; -.		
DR	PIR; B96750; B96750.		
DR	GO; GO:0005634; C:nucleus; IEA.		
DR	GO; GO:0003700; P:transcription factor activity; IEA.		
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.		
DR	InterPro; IPR001471; TF_ERF.		
DR	Pfam; PF00847; AP2-domain; 2.		
DR	PRINTS; PR00367; ETRSRSELMNT.		
DR	ProDom; PD001423; TF_ERF; 2.		
DR	SMART; SM00380; AP2; 2.		
SQ	SEQUENCE 425 AA; 48205 MW; C191BAS2D6C5ACBB CRC64;		
Query Match 29.7%; Score 926; DB 10; Length 425;			
Best Local Similarity 45.0%; Pred. No. 7.5e-55;			
Matches 207; Conservative 47; Mismatches 86; Indels 120; Gaps 12;			
QY	5	WLGFSLSP-----YEONHRRKDVYVSTTTTVDVAGEYCYDPTAASDESSAIQTSFPSP	58
DB	4	WLGFSLTPPLRICNSEEEELRHGSDGTL-----	32
QY	59	FGVVDAFTRDNNSHSRDW--DINGCACNNTHNDQDQPKLENFLGRT-----	104
DB	33	-----FFTH-----HPPVWRYDIN-----FDHHHDEDEKVEDLLSNHSQTYPINHNQT	78
QY	105	-----TTIYNTNENVG-----DGSQSGCYGGGSGGSLGLSMIKTLWLK-	143
DB	79	NNCTTVNRLNPPGVLHDDQVTPHYPLNPNLSNDYGGFERVG---SVSVFKSWLEQ	135
QY	144	-----NQPDVNDNQENGAAGK--LSLSNMNSTSCDNNDSNNV-	182
DB	136	GFAPFLSHYVTEAGTGNISHSNEETGYNTGSMLSLALSHGACDLSNESVSAR	195

```
QY 183 VAGKTIIDS-----VEATPKTIESFGORTSYRGVTRHRTGRYEAHLWDNSC 232
Db 196 VEEPVKDEKRLVVKPOKESVPRKSVDSYGQTSQYRGVTRHRTGRYEAHLWDNSC 255
QY 233 KREGQTRKGRQVYLGGYDKKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKHMTR 292
Db 256 KREGQTRKGRQAKIGGYDEEKAARAYDLAALKYWGPTTHLNFPLSNYEKEIEELNNMR 315
QY 293 QEYVASLRKSSGFSRGASIVGVTTRHHQHGHGWQARIQGVAGNKDLYLGTGTQEEAAEA 352
Db 316 QEFVAMLRNSSGFSRGASVYGVTRHHQHGHGWQARIQGVAGNKDLYLGTGTQEEAAEA 375
QY 353 YDIAAIKFRGLTAVTFMNMNRYNVKAILSPSLPIGSAK 392
Db 376 YDIAAIKFRGLNAVTFNDINRYDVKRICSSSTIVSDQAK 415
```

Search completed: July 15, 2004, 10:03:10
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 09:51:59 ; Search time 57 seconds
(without alignments)
2970.088 Million cell updates/sec

Title: us-09-980-364-2

Perfect score: 3119

Sequence: 1 MNNNLGFLSLSPYEQNHRRK.....YYHGGGGEVAPFTTWNNDN 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3119	100.0	579	4 AAB50873	Aab50873 Brassica
2	3119	100.0	579	4 AAB50872	Aab50872 Brassica
3	1310	42.0	707	6 ABR40766	Abr40766 Glycine m
4	1296	41.6	710	6 ABR40768	Abr40768 Glycine m
5	1131	36.3	692	6 ABR40877	Abr40877 Oryza sat
6	1119.5	35.9	553	6 ABR40771	Abr40771 Glycine m
7	1075.5	34.5	486	6 ABR40767	Abr40767 Glycine m
8	1047.5	33.6	574	7 ADD30879	Add30879 Plant yie
9	1047.5	33.6	574	7 ADD30334	Add30334 Plant yie
10	1047.5	33.6	574	7 ADD30823	Add30823 Plant yie
11	1040.5	33.4	516	6 ABR40836	Abr40836 Arabidops
12	1011	32.4	585	5 ABB79640	Abb79640 Cotton Al
13	1010.5	32.4	538	6 ABR40854	Abr40854 Gossypium
14	1003	32.2	470	6 ABR40853	Abr40853 Catalpa s
15	999	32.0	512	6 ABR40852	Abr40852 Glycine m
16	997.5	32.0	540	6 ABR40837	Abr40837 Arabidops
17	997	32.0	566	6 ABR40876	Abr40876 Arabidops
18	994	31.9	557	6 ABR40856	Abr40856 Glycine m
19	986	31.6	510	6 ABR40764	Abr40764 Glycine m
20	977	31.3	558	7 ADD30981	Add30981 Plant yie
21	974.5	31.2	498	7 ADD30350	Add30350 Plant yie
22	971	31.1	530	6 ABR40769	Abr40769 Glycine m
23	964.5	30.9	469	6 ABR40759	Abr40759 Oryza sat
24	961	30.8	492	6 ABR40855	Abr40855 Zea mays
25	957.5	30.7	384	6 ABR40770	Abr40770 Glycine m

ALIGNMENTS

RESULT 1

AAB50873
ID AAB50873 standard; protein; 579 AA.

XX AC AAB50873;

XX DT 19-MAR-2001 (first entry)

XX DE Brassica napus BNM3B protein.

XX KW Brassica napus microspore embryo; BNM; microspore embryogenesis; BNM3B; asexually derived embryo production; increased regenerative capacity.

XX OS Brassica napus.

XX PN EP1057891-Al.

XX PD 06-DEC-2000.

XX PF 02-JUN-1999; 99EP-00201745.

XX PR 02-JUN-1999; 99EP-00201745.

XX (CPRO-) CPRO-DLO CENT PLANTENVEREDELINGS REPROD.
(MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.

XX PI Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J; Boutillier K;

XX WPI; 2001-026720/04.

XX DR N-PSDB; AAC91399.

XX PT Brassica napus BNM3 DNA sequences, useful for transforming host cells to produce asexually derived embryos.

XX PS Claim 62; Fig 3; 51pp; English.

XX CC The present sequence is given in a specification relating to a Brassica napus microspore embryo (BNM) gene obtained during the induction of a microspore embryogenesis. The isolated DNA molecule can be used in a vector that is used to transform plant cells. The vector can be used in methods to produce asexually derived embryos, for modifying the regenerative capacity of a plant, and for producing an apomictic plant

XX SQ Sequence 579 AA;

Query Match 100.0%; Score 3119; DB 4; Length 579;

Best Local Similarity 100.0%; Pred. No. 3e-257;

26	951.5	30.5	642	5	ABR79638	Abb79638 Rice AINT
27	949.5	30.4	484	6	ABR40757	Abr40757 Zea mays
28	943.5	30.3	489	6	ABR40756	Abr40756 Zea mays
29	942	30.2	663	5	ABR79636	Abb79636 Soybean A
30	942	30.2	669	5	ABR79639	Abb79639 Rice AINT
31	941	30.2	415	7	ADD30560	Add30560 Plant yie
32	930	29.8	485	6	ABR40829	Abr40829 Zea mays
33	930	29.8	485	6	ABR40834	Abr40834 Zea mays
34	926	29.7	425	6	ABR40833	Abr40833 Arabidops
35	924	29.6	665	5	ABR79637	Abb79637 Soybean A
36	913.5	29.3	548	3	AAB07725	Aab07725 A Brassic
37	891	28.6	555	3	AAB07724	Aab07724 An Arabid
38	891	28.6	555	3	AAG31380	Aag31380 Arabidops
39	891	28.6	555	6	ABR40830	Abr40830 Arabidops
40	889	28.5	555	6	ABR40832	Abr40832 Arabidops
41	880	28.2	528	3	AAG31382	Aag31382 Arabidops
42	880	28.2	529	3	AAG31381	Aag31381 Arabidops
43	714	22.9	312	6	ABR40762	Abr40762 Glycine m
44	711.5	22.8	409	6	ABR40849	Abr40849 Momordica
45	681	21.8	345	6	ABR40875	Abr40875 Arabidops

Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNWLGSLSPYEQNHHRKDVYSSTTTTVDVAGEYCYDPTAADESSAQTSPSPFG 60
 Db 1 MNNWLGSLSPYEQNHHRKDVYSSTTTTVDVAGEYCYDPTAADESSAQTSPSPFG 60

QY 61 VVDAFTDRNNSHSDWDINGCACNNIHNDQDGPKEFLGRITTIYNTNENVGDGSGS 120
 Db 61 VVDAFTDRNNSHSDWDINGCACNNIHNDQDGPKEFLGRITTIYNTNENVGDGSGS 120

QY 121 GCGGGGGGSLGLSMKTLRNQPDVNDQENGAAGLISLMSNSTSCDNNNSNN 180
 Db 121 GCGGGGGGSLGLSMKTLRNQPDVNDQENGAAGLISLMSNSTSCDNNNSNN 180

QY 181 NVVAQKTIIDDSVEATPKKTIESFQRTSIYRGVTRHWTGRYEHLWDNSCKREGQTRK 240
 Db 181 NVVAQKTIIDDSVEATPKKTIESFQRTSIYRGVTRHWTGRYEHLWDNSCKREGQTRK 240

QY 241 GRQVILGGYDKEEKAARAYDLAALKYWGTTTTTNPMSYEYKEVEEMKHMTRQEVVASLR 300
 Db 241 GRQVILGGYDKEEKAARAYDLAALKYWGTTTTTNPMSYEYKEVEEMKHMTRQEVVASLR 300

QY 301 RKSSGFSRGASIRYGVTRHWHQGRWQARIGRVAGNKDLYLGTFTGTQEEAAAYDIAAKF 360
 Db 301 RKSSGFSRGASIRYGVTRHWHQGRWQARIGRVAGNKDLYLGTFTGTQEEAAAYDIAAKF 360

QY 361 RGLTAVTNFDMNRNVKAILSPSLPIGSAARLKEANRPVPSMMISNNVSESENSASG 420
 Db 361 RGLTAVTNFDMNRNVKAILSPSLPIGSAARLKEANRPVPSMMISNNVSESENSASG 420

QY 421 WQNAVQHHQGVLDLSLLHQHQRVNGYNGNLSSESARACFKQEDDQHHFLSNTQSLM 480
 Db 421 WQNAVQHHQGVLDLSLLHQHQRVNGYNGNLSSESARACFKQEDDQHHFLSNTQSLM 480

QY 481 TNIDHQSSVSDSVTCGVNGYQGYQGFAPVNCDAYAASEFDYNARNHYFFAQOQQTQ 540
 Db 481 TNIDHQSSVSDSVTCGVNGYQGYQGFAPVNCDAYAASEFDYNARNHYFFAQOQQTQ 540

QY 541 QSPGGDFPAAATNNVGSNNMYHGGGGEVAPFTTWNNDN 579
 Db 541 QSPGGDFPAAATNNVGSNNMYHGGGGEVAPFTTWNNDN 579

RESULT 2
 AAB50872
 ID AAB50872 standard; protein; 579 AA.
 AC AAB50872;
 DT 19-MAR-2001 (first entry)
 DE Brassica napus BNM3A protein.
 KW Brassica napus microspore embryo; BNM; microspore embryogenesis; BNM3A;
 KW asexually derived embryo production; increased regenerative capacity.
 OS Brassica napus.
 PN EP1057891-A1.
 PD 06-DEC-2000.
 PP 02-JUN-1999; 99EP-00201745.
 PR 02-JUN-1999; 99EP-00201745.
 PA (CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
 PE (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
 PI Van Lookeren Campagne M, Custers J, Miki B, Ouellet T, Hattori J;
 PI Boutillier K;
 XX WPI; 2001-026720/04.

DR N-PSDB; AAC91398.
 XX Brassica napus BNM3 DNA sequences, useful for transforming host cells to
 PT produce asexually derived embryos.
 XX Claim 61; Fig 3; slpp; English.
 CC The present sequence is given in a specification relating to a Brassica
 CC napus microspore embryo (BNM) gene obtained during the induction of
 CC microspore embryogenesis. The isolated DNA molecule can be used in a
 CC vector that is used to transform plant cells. The vector can be used in
 CC methods to produce asexually derived embryos, for modifying the
 CC regenerative capacity of a plant, and for producing an apomictic plant
 XX Sequence 579 AA;

Query Match 100.0%; Score 3119; DB 4; Length 579;
 Best Local Similarity 100.0%; Pred. No. 3e-257;
 Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNNWLGSLSPYEQNHHRKDVYSSTTTTVDVAGEYCYDPTAADESSAQTSPSPFG 60
 Db 1 MNNWLGSLSPYEQNHHRKDVYSSTTTTVDVAGEYCYDPTAADESSAQTSPSPFG 60

QY 61 VVDAFTDRNNSHSDWDINGCACNNIHNDQDGPKEFLGRITTIYNTNENVGDGSGS 120
 Db 61 VVDAFTDRNNSHSDWDINGCACNNIHNDQDGPKEFLGRITTIYNTNENVGDGSGS 120

QY 121 GCGGGGGGSLGLSMKTLRNQPDVNDQENGAAGLISLMSNSTSCDNNNSNN 180
 Db 121 GCGGGGGGSLGLSMKTLRNQPDVNDQENGAAGLISLMSNSTSCDNNNSNN 180

QY 181 NVVAQKTIIDDSVEATPKKTIESFQRTSIYRGVTRHWTGRYEHLWDNSCKREGQTRK 240
 Db 181 NVVAQKTIIDDSVEATPKKTIESFQRTSIYRGVTRHWTGRYEHLWDNSCKREGQTRK 240

QY 241 GRQVILGGYDKEEKAARAYDLAALKYWGTTTTTNPMSYEYKEVEEMKHMTRQEVVASLR 300
 Db 241 GRQVILGGYDKEEKAARAYDLAALKYWGTTTTTNPMSYEYKEVEEMKHMTRQEVVASLR 300

QY 301 RKSSGFSRGASIRYGVTRHWHQGRWQARIGRVAGNKDLYLGTFTGTQEEAAAYDIAAKF 360
 Db 301 RKSSGFSRGASIRYGVTRHWHQGRWQARIGRVAGNKDLYLGTFTGTQEEAAAYDIAAKF 360

QY 361 RGLTAVTNFDMNRNVKAILSPSLPIGSAARLKEANRPVPSMMISNNVSESENSASG 420
 Db 361 RGLTAVTNFDMNRNVKAILSPSLPIGSAARLKEANRPVPSMMISNNVSESENSASG 420

QY 421 WQNAVQHHQGVLDLSLLHQHQRVNGYNGNLSSESARACFKQEDDQHHFLSNTQSLM 480
 Db 421 WQNAVQHHQGVLDLSLLHQHQRVNGYNGNLSSESARACFKQEDDQHHFLSNTQSLM 480

QY 481 TNIDHQSSVSDSVTCGVNGYQGYQGFAPVNCDAYAASEFDYNARNHYFFAQOQQTQ 540
 Db 481 TNIDHQSSVSDSVTCGVNGYQGYQGFAPVNCDAYAASEFDYNARNHYFFAQOQQTQ 540

QY 541 QSPGGDFPAAATNNVGSNNMYHGGGGEVAPFTTWNNDN 579
 Db 541 QSPGGDFPAAATNNVGSNNMYHGGGGEVAPFTTWNNDN 579

RESULT 3
 ABR40766
 ID ABR40766 standard; protein; 707 AA.
 AC ABR40766;
 DT 16-MAY-2003 (first entry)
 DE Glycine max oil trait related protein sequence SEQ ID NO:352.
 KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW

KW Lip15-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CKC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant.
 XX Glycine max.
 XX WO2003002751-A2.
 XX 09-JAN-2003.
 XX 27-JUN-2002; 2002WO-US020152.
 XX 29-JUN-2001; 2001US-0301913P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;
 XX WPI; 2003-201509/19.
 XX N-PSDB; ACC00801.
 XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX Claim 12; Page 369-371; 542pp; English.
 XX The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, Lip15-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity and CKC-like transcription factor activity. Also described: (1)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (II) can also be
 CC used as a hybridisation probe. ACC00826 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX Sequence 707 AA;
 SQ
 Query Match 42.0%; Score 1310; DB 6; Length 707;
 Best Local Similarity 45.6%; Pred. No. 1.4e-102;
 Matches 322; Conservative 77; Mismatches 139; Indels 168; Gaps 29;
 QY 4 NLLGFLSPYEQNHRRKDVYSST-----TTVDVAGEYCVDPDTAASDESSAIQ 52
 DB 5 NLLGFLSPHEEPSSQD-HSQITPSPFPNPGSISSTDVAGG-CFDJLTS---DSTPHL 59
 QY 53 TSPSPFPGVVDAFTDNN-SHGRDMDIN-----GCACN--NHNDEQDGKLENFL 101
 DB 60 LNLFS-YG-IYEAFHRNNSINTQDWKENYNSQNLGLTSCNQNNQNNQQQKLENFL 117
 QY 102 GRTTITNTNENV-GDGGSGGCY-----GGGDGGG-----GSLGLSMKTLWR 143
 DB 118 GGHG--FGEHQYVGNASTDWFPAQPVAGGSGGSGGNNNNNSNLSGLSNKTLWR 175
 QY 144 NOPVDNVNONGNAKG-----LSLSMN-----SST-----SCDNNDSNN-- 181
 DB 176 NQP-PNSENINNNNSCGNIRSSVQQTLSLSMGTGSSQSTSLPLLTAASVDNGSSSDNQ 234
 QY 182 ---VVAQGIIDSVEATPKKTIESTFQRTSIYRGVTRHWTGCTGYEAHLWDNSCKREGOT 238
 DB 235 PNTSAALDSTQTGAIEFAPRKSIDTFQRTSIYRGVTRHWTGCTGYEAHLWDNSCKREGOT 294

QY 239 RKGRQVILGGYDKEEKAARAYDLAALKYWGCTTTTTFPMSEYEXEVEEMKMTQEVYAS 298
 DB 295 RKGRQVILGGYDKEEKAARAYDLAALKYWGCTTTTTFPISEYKELEBEMKMTQEVYAS 354
 QY 299 LRRKSSGFSRGASIVRGVTRHHQHGRWQARIGRVAGNKDLVLTGTFGTQEEAAEAYDAIAI 358
 DB 355 LRRKSSGFSRGASIVRGVTRHHQHGRWQARIGRVAGNKDLVLTGTFGTQEEAAEAYDAIAI 414
 QY 359 KFRGLTAVTNFDMRYNVKAILLESPLSIGSAARLKE-----ANRPVPSMM 405
 DB 415 KFRGLSAYTNFDMRGYDVKSILESTTLPFGAAARLKDMEQVELSDVNGHRAADQVDHSII 474
 QY 406 MISNNVSSSENSASGWAQAAVQHGHQVDLSLLHGHQERYNGYYNGGNLSSESARACFKQ 465
 DB 475 MSSHLTQGINNNYAGGTAT--HHWNHNAHAFHQPQCTTWHYFYGQRIN-----WCKQE 527
 QY 466 EDDQ-----HHFLSNTQSL--MTNIDHQSSVSDSDSVTVCGNVV 501
 DB 528 QODNSDAPHSLSYSDIHQQLGNGNTHNFHTNSGLHFMLSMD---SASIDNSSSSNSV 584
 QY 502 --GYGGYOGF-AAPV-NCDAYAASEFDVN-----SASVDTVKASAYDQGSACNTWVPT 679
 DB 585 YDYGCGGGGYNMPTGTTTAVVAGSDGONPRSNHGFQDNEIKALGYESVYGSATDSYHAH 644
 QY 527 ARNHYYFAQQOQTQOSPGGDFPAAAMNNVGNMYYHGGEGGEVAPT 572
 DB 645 ARNLYYLTQQQ-----SSSVDTVKASAYDQGSACNTWVPT 679

RESULT 4

ABR40768
 ID ABR40768 standard; protein; 710 AA.
 XX ABR40768;
 AC ABR40768;
 DT 16-MAY-2003 (first entry)
 XX Glycine max oil trait related protein sequence SEQ ID NO:356.
 DE Glycine max oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW Lip15-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CKC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant.
 XX Glycine max.
 XX WO2003002751-A2.
 XX 09-JAN-2003.
 XX 27-JUN-2002; 2002WO-US020152.
 XX 29-JUN-2001; 2001US-0301913P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 PI Tarczynski MC;
 XX WPI; 2003-201509/19.
 XX N-PSDB; ACC00803.
 XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX Claim 12; Page 374-376; 542pp; English.
 XX The present invention describes an isolated nucleotide fragment (I)
 CC

comprising a nucleic acid sequence (NS) chosen from a NS encoding a polypeptide (PP) having receptor-like protein kinase activity, mitogen activated protein (MAP)-kinase activity, LIP15-like transcription factor activity, calceosin-like activity, ATP citrate lyase activity, SNF1-like activity, and CKC-like transcription factor activity. Also described: (I) complement (II) of (I); (2) a chimeric construct (III) comprising (I) or (II), operably linked to a regulatory sequence; (3) a plant (IV) comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5) oil obtained from (V). (I) or its part can be used in antisense inhibition or co-suppression in a transformed plant. (III) is useful for altering the oil phenotype in a plant such as corn, soybean, wheat, rice, canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for creating transgenic plants having altered lipid profiles. (I) can also be used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to ABR40879 represent sequences used in the exemplification of the present invention

XX
SQ Sequence 710 AA;

Query Match 41.6%; Score 1296; DB 6; Length 710;
Best Local Similarity 45.5%; Pred. No. 2.3e-101;
Matches 324; Conservative 66; Mismatches 146; Indels 176; Gaps 29;

QY 4 NWLGSLSPYE-----QNHR-----KDYSTTTTVDVAGYCYDPTAASDESSA 50
DB 5 NLLGSLSPQHPSSQDHSQTAPRCFNPDISS-----DVAGD-CFGLTS-----DSTP 55
QY 51 IQTSPSPFGVVDAFTDRDNNSH-SRDWDIN-----GCACNN-----IHN-----DEQDGP 95
DB 56 HLLNLP-S-YG-IYFAFHRSHNIHTQDKENYNQNLGTCSCNQNNHNNHQOQQQP 113
QY 96 KLENFLGRTTIYNTNEN-VDGSGSGCY-----GGDGG-----GSLGLSM 137
DB 114 KLENFLGHS--FGEHQPGYCGNSASTYMPAPQVLAGGGGGGSGNSSTNSSSIGLSM 171
QY 138 IKTWLRNPQVDNDQNGNAAG-----LSLSNN-----SST-----SCDNNND 177
DB 172 IKTWLRNPQPHSENNNNNSGNSRVSQVTLISLMSGTSQSTSLPLLTASVDNGES 231
QY 178 SNN-----VVAQGTIDDSVEATPKTIESFGORTSIYRGVTRHRTWGYEAHLWNSC 232
DB 232 SSDNKQPHTTAALDTQTGA-ETAPRKSIDTFGORTSIYRGVTRHRTWGYEAHLWNSC 291
QY 233 KREGTRKRGVYLGQYDKKKAARAYDLAALKYWGTTTTTNPMSYKEVEEMKMTNR 292
DB 292 RREGTRKRGVYLGQYDKKKAARAYDLAALKYWGTTTTTNPFIHYKELEEMKMTNR 351
QY 293 QEYVASLRKSGFSGASIRYGVTRHHQGRWQARIGRVAGNKDLYLGTFTQEEAAEA 352
DB 352 QEYVASLRKSGFSGASIRYGVTRHHQGRWQARIGRVAGNKDLYLGTFTQEEAAEA 411
QY 353 YDIAIKRGLTAVTNEDPMYNNKATLESPLPIGSAKELKE-----ANR 399
DB 412 YDVAIKRGLSAVTNPMYNDYVKSILESTLPIGGAKEKDKMEQVELRVENHRAQ 471
QY 400 PVPMMTISNNVSESENASQWQAAGVGHQGVLDLSLHQHQRINGYYNGNLSSES 459
DB 472 EDHSSIMNSHLTQGIINNYAAGTTATHHNNHNAFAHQPPCTTIHYPGQRIN--- 527
QY 460 RACFKQEDDQ-----HFLSLTQSL--MTNIDHQSVDSDSVT 495
DB 528 --WCKQEQDNDASHLSYSIDHQLQGNNTNFTNTHFHTNSGLHPLMSND---SASIDNS 582
QY 496 VCGNVV--GYGQYQGF-AAPV-NCDAVAASEFDYN----- 526
DB 583 SSNSVYDYGGGGGYNNVPIWGTTTTTVDVANDQNPNSRHGFGDNEIKALGYESVYSGTT 642
QY 527 -----ARNHYFAQQQTQSPGDPFAAMTNVNVGSMYVHGGGGEVAPT 572
DB 643 DPYHAHAENLYLTQQQPS-----VDAAKASAYDQGSACNTWVPT 683

RESULT 5

ABR40877
ID ABR40877 standard; protein; 692 AA.
AC ABR40877;
DT 16-MAY-2003 (first entry)
XX Oryza sativa oil trait related protein sequence SEQ ID NO:530.
DE Oryza sativa oil trait related protein sequence SEQ ID NO:530.
KW plant; oil trait; oil phenotype; altered lipid profile; MAP kinase; receptor-like protein kinase; mitogen activated protein kinase; oil; LIP15-like transcription factor calceosin; ATP citrate lyase; SNF1; CKC-like transcription factor; antisense inhibition; co-suppression; transgenic plant.
XX Oryza sativa.
XX WO2003002751-A2.
XX 09-JAN-2003.
XX 27-JUN-2002; 2002WO-US020152.
XX 29-JUN-2001; 2001US-03019139.
XX (DUPO) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT, Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B, Tarczynski MC;
XX WPI; 2003-201509/19.
XX Novel nucleotide fragment encoding polypeptides having receptor-like protein kinase activity, calceosin-like activity, useful for altering oil phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
XX Claim 12; Page 538-540; 542pp; English.
XX The present invention describes an isolated nucleotide fragment (I) comprising a nucleic acid sequence (NS) chosen from a NS encoding a polypeptide (PP) having receptor-like protein kinase activity, mitogen activated protein (MAP)-kinase activity, LIP15-like transcription factor activity, calceosin-like activity, ATP citrate lyase activity, SNF1-like activity, and CKC-like transcription factor activity. Also described: (1) complement (II) of (I); (2) a chimeric construct (III) comprising (I) or (II), operably linked to a regulatory sequence; (3) a plant (IV) comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5) oil obtained from (V). (I) or its part can be used in antisense inhibition or co-suppression in a transformed plant. (III) is useful for altering the oil phenotype in a plant such as corn, soybean, wheat, rice, canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for creating transgenic plants having altered lipid profiles. (I) can also be used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to ABR40879 represent sequences used in the exemplification of the present invention

XX
SQ Sequence 692 AA;

Query Match 36.3%; Score 1131; DB 6; Length 692;
Best Local Similarity 40.1%; Pred. No. 2.7e-87;
Matches 278; Conservative 65; Mismatches 125; Indels 236; Gaps 23;

QY 3 NNWLGFSLSPYQNHRRKDVYSSSTTTTVDVAGYCYDPTAASDESSA IQTSPFGV 62
DB 5 NNWLAFSLSPQQLPP-----SQNSTLSIAA-----TTTAGDSSGDCVFNIP----- 50
QY 63 VDAFTDRDNNSHRDWDINGCACNNTNDEQGPKEFLGRTTIYNTNENVG----- 117
DB 51 -----QDWSMRGSELSALVAE-----PKLEDFLGISFSEQQHHGGKGVIPS 94
QY 118 SGSGCY-----GGG-----DG 128

Db 95 SAAACVASSGSSVGLYPPSSSSLOPADSVMTATSSPVVAHDGVSGGWSVAAAAAAS 154
QY 129 GGSGLGSMKTLWLRNOPVDNVQENGAAGLSLMSNS----- 168
Db 155 GNGGIGLSMKWLRSQFAP-----QPAQALSLSMNMAGTTTAQGGGAMALLAGAGE 206
QY 169 -----STSC-----DNNNDNNNVAAQKTIDDS-----V 193
Db 207 RGRTPPASELSITSAGCATATTWAGGKEINEGSGSAGAVAVGSESGSNAVVEAGAA 266
QY 194 EATPKKTIIEFGORTSYRGVTRHRTWGRVEAHLWNSCKRGQTRKGRQVYLGVDKKE 253
Db 267 AAARKSVDFGORTSYRGVTRHRTWGRVEAHLWNSCKRGQTRKGRQ--GGVDKKE 323
QY 254 KAARAYDLAALKYGGTTTTFNPFMSVEKEVEEMKMTROEYVASLRKSSGFSRGASTY 313
Db 324 KAARAYDLAALKYGGPTTTFNPFVNNYEKELEEMKMTROEYVASLRKSSGFSRGASTY 383
QY 314 RGVTTRHHQGRWQARIGRVAGNKDLYLGTFGTQEEAAEAYDIAAIFRGLTAVTNFDMNR 373
Db 384 RGVTTRHHQGRWQARIGRVAGNKDLYLGTFGTQEEAAEAYDIAAIFRGLNAVTFNFMGR 443
QY 374 YNVKAILLESPLTGSAAKELKANRVPV-SMMWISNNVS-----ESSENSASGW 421
Db 444 YVKSILDSALPVGTAARKLKDAEAAAAYDVGRVIAHLGGDGAYAAHVGHHHHSAAAAW 503
QY 422 QNAAVQ-----HHQGVDLSSLHQQEYRGYNYNGGNLSSESARACFKQEDDQHHFLS 474
Db 504 PTIAFAAAAAPPHPAAG---LYHPYAQPLRGW-----C-KQEQD-HAVTA 543
QY 475 NTQSLTNIDHQSVDSDSVTCGVVGVGGYOGPAAPVNCDAYAASEFDYARNHYIFA 534
Db 544 AHSLO-DLHH-----LNLGAAAAA-----HDFFS 567
QY 535 QQOQTQSGCGDPPAAMTNVNGSN-MYIHGEGG 567
Db 568 QAMQOQHGLSIDNASLEHSTGNSVYNGDNGG 601

RESULT 6
ABR40771
ID ABR40771 standard; protein; 553 AA.
AC ABR40771;
XX 16-MAY-2003 (first entry)
DT Glycine max oil trait related protein sequence SEQ ID NO:362.
DE Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
KW receptor-like protein kinase; mitogen activated protein kinase; oil;
KW Lip15-like transcription factor caleosin; ATP citrate lyase; SNF1;
KW CKC-like transcription factor; antisense inhibition; co-suppression;
transgenic plant.
XX Glycine max.
OS
XX WO2003002751-A2.
FN
XX 09-JAN-2003.
PD
XX 27-JUN-2002; 2002WO-US020152.
PF
XX 29-JUN-2001; 2001US-0301913P.
PR
XX (DUPO) DU PONT DE NEMOURS & CO E I.
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
PI Tarczynski NC;

DR WPI: 2003-201509/19.
XX N-PSDB; ACC00806.
PT Novel nucleotide fragment encoding polypeptides having receptor-like
PT protein kinase activity, caleosin-like activity, useful for altering oil
PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
XX
PS Claim 12; Page 382-383; 542pp; English.
XX
CC The present invention describes an isolated nucleotide fragment (I)
CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
CC activated protein (MAP)-kinase activity, Lip15-like transcription factor
CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
CC activity, and CKC-like transcription factor activity. Also described: (1)
CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
CC (II), operably linked to a regulatory sequence; (3) a plant (IV); and (5)
CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and
CC oil obtained from (V). (I) or its part can be used in antisense
CC inhibition or co-suppression in a transformed plant. (III) is useful for
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
CC canola, Brassica, sorghum, sunflower or coconut. (II) is also useful for
CC creating transgenic plants having altered lipid profiles. (I) can also be
CC used as a hybridisation probe. ACC00806 to ACC00868 and ABR40591 to
CC ABR40879 represent sequences used in the exemplification of the present
XX invention
SQ Sequence 553 AA;
Query Match 35.9%; Score 1119.5; DB 6; Length 553;
Best Local Similarity 43.5%; Pred. No. 1.8e-86;
Matches 275; Conservative 77; Mismatches 147; Indels 133; Gaps 23;
QY 1 MNWNWGLSGLSPYEQNHHRKDVSSSTTTTVDVAGEYCYDPTAASDESSAQTGFPSPFG 60
Db 1 MNWNWLSFPLS-----FTHSSLPAHDLQATQYHQS 31
QY 61 VVDAFTDNKSHSRDWD-INGCACNNHNDQGPKEFLGRTTITNTYNNVGDGSG 119
Db 32 LGLVNNNDNPNFQNHDMNLINTHSNEI-----PKVADFLG-----VSKSENQSDLAA 79
QY 120 SGCYGGGSGGSGSLGSLKTLWLRNOPVDNVN--QENGNA-AKGLSLMSNS--STSCDN 174
Db 80 LNEIHSNDSYLFNNSLVP--MQNPVLDTPSNYQENANSLQSLTSLMSGKSTCET 137
QY 175 NDSNNNVAAOGKTIDDSVEATPKKTIIEFGORTSYRGVTRHRTWGRVEAHLWNSCKR 234
Db 138 SCENSTNT-----TVEVAPRRTLDTFGORTSYRGVTRHRTWGRVEAHLWNSCKR 188
QY 235 EGOTRKGQVYLGVDKKEKAAAYDLAALKYGGTTTTFNPFMSVEKEVEEMKMTROE 294
Db 189 EGOSRKGQVYLGVDKKEKAAAYDLAALKYGGTTTTFNPFMSVEKEVEEMKMTROE 248
QY 295 YVASLRKSSGFSRGASTYRGVTRHRTWGRWQARIGRVAGNKDLYLGTFGTQEEAAEAYD 354
Db 249 FVAALRRKSSGFSRGASMYRGVTRHRTWGRWQARIGRVAGNKDLYLGTFGTQEEAAEAYD 308
QY 355 IAAIKERGLTAVTNFDMNRVNNYKAILLESPLPI-GSAAKRLKEA-----NRPVSMMLI- 407
Db 309 IAAIKERGLNAVTFNPFMSRYDVVKAILLESNTLPIGGGAARLKEAQALESKRREMTALG 368
QY 408 SNNVSESENSASGWNAA---VQHH-----QGVDLSLLH-QHQQ---RYNGY 448
Db 369 SSSITFYGTSSASSRLHAYPLMQHHHQEQPQLLTQNHDISSSHFSHQDPLHHQYI 428
QY 449 YNGNLSSSARACFK-QEDDQHH--FLSNTQSL---MTNIDHSSVS-----DVSVTVC 497
Db 429 QTQLQLHQOQSGASSYFQNNAAQFYNGLYQHPALQGMNMNMGSSSSSSSVLENNNSNN 488
QY 498 GNVVYGGYQGFAPAVNCDA-----YAASEFDYARNHYIFAQQOQTQSGPD 546
Db 489 NVVGFGVG-SGFGMASNATAGTNTVGTAEELGLVKVDYDMPAGGCGGWAADSMQTSNGG- 546

QY 547 FPAAMTNNVGSNNYYHGBGGGEVAPTFTVND 578

DB 547 -----VFTMND 553

RESULT 7

ABR40767
ID ABR40767 standard; protein; 486 AA.

XX ABR40767;

DT 16-MAY-2003 (first entry)

DE Glycine max oil trait related protein sequence SEQ ID NO:354.

XX Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
KW receptor-like protein kinase; mitogen activated protein kinase; oil;
KW LIP15-like transcription factor caleosin; ATP citrate lyase; SNF1;
KW CKC-like transcription factor; antisense inhibition; co-suppression;
transgenic plant.

XX Glycine max.

XX WO2003002751-A2.

PN 09-JAN-2003.

PD 27-JUN-2002; 2002WO-US020152.

PF 29-JUN-2001; 2001US-0301913P.

PR (DUPO) DU PONT DE NEMOURS & CO E I.

PA (PION-) PIONEER HI-BRED INT INC.

XX Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
PI Jones TJ, Kinney AJ, Klein TW, Li C, Oliveira IC, Sakai H, Shen B;
PI Tarczynski MC;

XX WPI; 2003-201509/19.

DR N-PSDB; ACC00802.

XX Novel nucleotide fragment encoding polypeptides having receptor-like
PT protein kinase activity, caleosin-like activity, useful for altering oil
PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.

PS Claim 12; Page 372-373; 542pp; English.

XX The present invention describes an isolated nucleotide fragment (I)
CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
CC polypeptide (pp) having receptor-like protein kinase activity, mitogen
CC activated protein (MAP)-kinase activity, LIP15-like transcription factor
CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
CC activity and CKC-like transcription factor activity. Also described: (I)
CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
CC oil obtained from (V). (I) or its part can be used in antisense
CC inhibition or co-suppression in a transformed plant. (III) is useful for
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
CC creating transgenic plants having altered lipid profiles. (I) can also be
CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
CC ABR40879 represent sequences used in the exemplification of the present
CC invention

SQ Sequence 486 AA;

Query Match 34.5%; Score 1075.5; DB 6; Length 486;
Best Local Similarity 50.2%; Pred. NO. 8.7e-83;
Matches 240; Conservative 45; Mismatches 99; Indels 95; Gaps 12;

QY 167 NSTSTCDNNNDNNVACGTTIDSVETAPKKTIESFGORTSIYRGVTRHRTGRVEAH 226

DB 4 NGESSSDNKQP--NTSAALDSTQTCAGIETAPRKSIDTFGORTSIYRGVTRHRTGRVEAH 61
QY 227 LWDNSCKREGQTRKGRQVYLGIDYKKEKAAAYDLAALKYAGTTTNTNPFMSYEKEVEE 286
DB 62 LWDNSCRREGQTRKGRQVYLGIDYKKEKAAAYDLAALKYAGTTTNTNPFISHEKELEE 121
QY 287 MCHMTROEYVASLRKSSGFGSRGASIRGVTRHHQHGWRQARIQVAGNKDLYLGTGTQ 346
DB 122 MCHMTROEYVASLRKSSGFGSRGASIRGVTRHHQHGWRQARIQVAGNKDLYLGTGTQ 181
QY 347 EBAEAYDIAIKFRGLTAVTNFOMNRYNVKALLESPLPIGSAAKRLKE----- 396
DB 182 EBAEAYDVAIKFRGLSAVTNFMRSRYDVYSIIESTITLPIGGAARLKDQEQVELSDVN 241
QY 397 ---ANRPVPSMMIMSNVSESENASGWCNAVQHHQGVDSLHLFHQHQERYNGYNGGN 453
DB 242 GHRADQVDHSIIIMSSHLTQGINNVAGGTAT--HNNWNAHAFQPOCTTMHYPYQOR 299
QY 454 LSSESARACFKQEDDQ-----HHFLSNTQSL--MTNIDHQSSV 489
DB 300 IN-----WCKQEQDQNSDAPHSLSYSDIHQQLQGNNGTHNFFHTNSGLHPMLSMD---SA 351
QY 490 SDDSVTVCGNVV--GYGGYQGF-AAAPV-NCDAAYAASEFDYN----- 526
DB 352 SIDNSSSSNVYDYGGGGGVNVMPGTTTAVVASDGDQNPERSNHGFGDNEIKALGYES 411
QY 527 -----ARNHYFAOQQQTQSPGDFPAAMTNNVGSNNYYHGBGGGEVAPT 572
DB 412 VYGSATDSYHAHARNLYLTQQQ-----SSVDTVKASAYDQGSACNTWVPT 458

RESULT 8

ADD30879
ID ADD30879 standard; protein; 574 AA.

XX AC ADD30879;

DT 15-JAN-2004 (first entry)

DE Plant yield-related protein from clone G1793.

XX transcription factor; transgenic plant; growth rate; senescence;
KW seed germination rate; plant vigor; seedling vigor.

XX Arabidopsis thaliana.

XX WO2003013227-A2.

XX 20-FEB-2003.

PF 09-AUG-2002; 2002WO-US025805.

XX 09-AUG-2001; 2001US-0310847P.

PR 19-NOV-2001; 2001US-0336049P.

PR 11-DEC-2001; 2001US-0338692P.

PR 14-JUN-2002; 2002US-00171468.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

XX Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;

PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;

PI Brown PE;

XX WPI; 2003-248221/24.

DR N-PSDB; ADD30878.

XX New plant transcription factor polynucleotides and polypeptides, useful
PT in producing transgenic plants with commercially valuable properties,
PT such as an alteration in a plant growth characteristic, e.g. growth rate
or apomixis.

XX Disclosure; SEQ ID NO 908; 454pp; English.

CC The invention relates to a number of isolated Arabidopsis thaliana cDNA
CC sequences and their encoded proteins which are especially transcription
CC factor related cDNA's and proteins. The isolated or recombinant plant
CC transcription factor polynucleotides and polypeptides are useful in
CC producing transgenic plants with commercially valuable properties, i.e.
CC modified or altered desirable traits as compared to a reference plant,
CC such as an alteration in a plant growth characteristic, e.g. growth rate,
CC germination rate of seeds, vigor of plants and seedlings, or leaf and
CC flower senescence. Sequence information related to the polynucleotides
CC and polypeptides can also be used in bioinformatic search methods. The
CC transgenic plant is useful for growing a progeny plant from a parent
CC plant. This sequence represents one of the proteins of the invention.
XX
SQ Sequence 574 AA;

Query Match 33.6%; Score 1047.5; DB 7; Length 574;
Best Local Similarity 38.7%; Pred. No. 2.8e-80;
Matches 256; Conservative 83; Mismatches 148; Indels 175; Gaps 20;

QY 2 NNNWLGPF-----SLSPYEQNHHRKDVYSSTTTVDVAGEYCYDPTAASDESSAIQIS 54
Db :|||||
3 SNNWLGPF.LSPNSSLPPHEYN-----

QY 55 FPSFGVVDVDAFTRDNNSHRDWDINGCACNNIHNDQDG-----PKLENFLGRITTIYNT 110
Db :|||||
25 -----LGLVSDHM--DNPFQTQENM-----INPHGGGDEGGEVVKVADFLGVSKPDENQ 73

QY 111 NENVGSGSGCYGGGGGSLGSMIKTWLRNQPV-----DNVDN-----QENGNA 158
Db :|||||
74 SNHLVAYNDSYFFHTN-----SLMPSVQSNVDVVAACDNTFNNSSYHELOESAHN 125

QY 159 AKGLSLSMNSTSCDNNNDNNNVV---AQGKTIDDSVE-----ATPKKTIESFQ 206
Db :|||||
126 LQSLTLLSMGTT-----AGNNVVDKASPSTTGDNASGGALAVVETATPRALDTFQ 177

QY 207 RTSYRGVTRHRTGTYEAHLWDNSCKREGQTRKGRQVYLGVDYDKEEKAARYDLAALKY 266
Db :|||||
178 RTSYRGVTRHRTGTYEAHLWDNSCKREGQTRKGRQVYLGVDYDKEEKAARYDLAALKY 237

QY 267 WGTITTFNPFMSYEKEVEEMKMTROEYVASLRKSSGFSRGASYRGVTRHHQHRWQ 326
Db :|||||
238 WGPSTTTFNPFITNYKEVEEMKMTROEFVAAIRKSSGFSRGASMTRGVTRHHQHRWQ 297

QY 327 ARIGRVAGNDLYLGTGTQEEAAYDIAAIFRGLTAVTNFDMRYNKAILESPLP 386
Db :|||||
298 ARIGRVAGNDLYLGTSTEEAAYDIAAIFRGLNAVTFEINRYDVKAILESSTLP 357

QY 387 I-GGAARLKEA-----
Db :|||||
358 IGGGAARLKEAQALESSRKREAEMLALGSSFOYGGSGSTGSGSTSRLLQLQYPYLSIQQ 417

QY 400 PVPSSMMISNNVSESENASGQWAAVOHQGVDSLHLHQHRYNGYNGNLSSES 459
Db :|||||
418 PLEFPLSLQNDISHYNNNNHDSSEFNHHSYIQ-TQLHLHQNTNVLQQSSQNSQQLY 476

QY 460 RACFQKQEDDQHFLSNTQSLMTNIDHQ--GSVSDSVTVCGNVVGYGQGAAPV-NCD 516
Db :|||||
477 NAYLHSPALLHGLVST-SIVDNNNNNGSGSYNTAAFLGN---HGIGIGSSSTVGSTE 532

QY 517 AYASEPDYNARNHYFAQQOQTQSGPDPPAAMTNNVGSNMYHGGEGEVAFTPTVW 576
Db :|||||
533 EFTPTKTDYD-----MPSSDGTGGYSGWTSVSQSPNGG-----VETMW 572

QY 577 ND 578
Db :
573 NE 574

RESULT 9
ADD30334
ID ADD30334
XX standard; protein; 574 AA.
AC ADD30334;

XX 15-JAN-2004 (first entry)
XX Plant yield-related protein from clone G1793.
XX
XX transcription factor; transgenic plant; growth rate; senescence;
KW seed germination rate; plant vigor; seedling vigor.
XX
XX Arabidopsis thaliana.
XX
XX WO2003013227-A2.
XX
XX 20-FEB-2003.
XX
XX 09-AUG-2002; 2002WO-US025805.
XX
XX 09-AUG-2001; 2001US-0310847P.
PR 19-NOV-2001; 2001US-0336049P.
PR 11-DEC-2001; 2001US-0338692P.
PR 14-JUN-2002; 2002US-00171468.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
XX Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;
PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pinada O, Yu G;
PI Broun PE;
XX
XX WPI: 2003-248221/24.
DR N-PSDB; ADD30333.
XX
XX New plant transcription factor polynucleotides and polypeptides, useful
PT in producing transgenic plants with commercially valuable properties,
PT such as an alteration in a plant growth characteristic, e.g. growth rate
PT or apomixis.
XX
XX Disclosure; SEQ ID NO 363; 454pp; English.
XX
XX The invention relates to a number of isolated Arabidopsis thaliana cDNA
CC sequences and their encoded proteins which are especially transcription
CC factor related cDNA's and proteins. The isolated or recombinant plant
CC transcription factor polynucleotides and polypeptides are useful in
CC producing transgenic plants with commercially valuable properties, i.e.
CC modified or altered desirable traits as compared to a reference plant,
CC such as an alteration in a plant growth characteristic, e.g. growth rate,
CC germination rate of seeds, vigor of plants and seedlings, or leaf and
CC flower senescence. Sequence information related to the polynucleotides
CC and polypeptides can also be used in bioinformatic search methods. The
CC transgenic plant is useful for growing a progeny plant from a parent
CC plant. This sequence represents one of the proteins of the invention.
XX
SQ Sequence 574 AA;

Query Match 33.6%; Score 1047.5; DB 7; Length 574;
Best Local Similarity 38.7%; Pred. No. 2.8e-80;
Matches 256; Conservative 83; Mismatches 148; Indels 175; Gaps 20;

QY 2 NNNWLGPF-----SLSPYEQNHHRKDVYSSTTTVDVAGEYCYDPTAASDESSAIQIS 54
Db :|||||
3 SNNWLGPF.LSPNSSLPPHEYN-----

QY 55 FPSFGVVDVDAFTRDNNSHRDWDINGCACNNIHNDQDG-----PKLENFLGRITTIYNT 110
Db :|||||
25 -----LGLVSDHM--DNPFQTQENM-----INPHGGGDEGGEVVKVADFLGVSKPDENQ 73

QY 111 NENVGSGSGCYGGGGGSLGSMIKTWLRNQPV-----DNVDN-----QENGNA 158
Db :|||||
74 SNHLVAYNDSYFFHTN-----SLMPSVQSNVDVVAACDNTFNNSSYHELOESAHN 125

QY 159 AKGLSLSMNSTSCDNNNDNNNVV---AQGKTIDDSVE-----ATPKKTIESFQ 206
Db :|||||
126 LQSLTLLSMGTT-----AGNNVVDKASPSTTGDNASGGALAVVETATPRALDTFQ 177

QY 207 RTSYRGVTRHRTGTYEAHLWDNSCKREGQTRKGRQVYLGVDYDKEEKAARYDLAALKY 266
Db :|||||

Db 178 RTSIYRGVTRHRTWGRYEHLWDNSCRREGQSKGRQVYLGSDYKEDKAARSYDLAALKY 237
 QY 267 WTTTTTNPMSGEYKEVEEMKHMTRQYVVASLRKSGFGRGASIRYGVTRHHQHGSEW 326
 Db 238 WGPSTTTNPITNYEKEVEEMKHMTRQEFVAAIRKSGFGRGASIRYGVTRHHQHGSEW 297
 QY 327 ARIGRVAGNKDLYLGTGCTGTEBAEAYDIAAKFGLTAFTNFDNRNRYNKAILESPL 386
 Db 298 ARIGRVAGNKDLYLGTGCTGTEBAEAYDIAAKFGLTAFTNFDNRNRYNKAILESPL 357
 QY 387 I-GSAAKRLKEA-----NR 399
 Db 358 IGGGAAKRLKEAQLSSRKREAEIALGSSFYQGGSGTSGTSSRLQIQPYPLSIQ 417
 QY 400 PVPMMISNNVSESENGSQNAVQHQVDLSLLHQHQRVNGYNGNLSSEA 459
 Db 418 PLEPFLSQNDISHYNNNAHDSSFNHSHYIQ-TQLHLQQTNNYVYVQSSQNSQOLY 476
 QY 460 RACFKQEDDQHFLSNTQSLMTNIDHQ--SSVSDSVTVCGNVYGYGQGAAPV-NCD 516
 Db 477 NAYLHSPALLHGLVST-SIVDNNNGSGSGSYNTAAFLGN--HGIGSSSTVGSTE 532
 QY 517 AYAASEFDYNARNHYFAQQOQTQSGDPFPAAMTNVGNMYHGGEGGEVAPTFTVW 576
 Db 533 EPTVTKTDYD-----MPSSDGTGGYSGWTSBVSQGSNPGG-----VFTMW 572
 QY 577 ND 578
 Db 573 NE 574

RESULT 10

ADD30823
 ID ADD30823 standard; protein; 574 AA.

XX AC ADD30823;

XX DT 15-JAN-2004 (first entry)

XX DE Plant yield-related protein from clone G1793.

XX KW transcription factor; transgenic plant; growth rate; senescence;
 KW seed germination rate; plant vigor; seedling vigor.

XX OS Arabidopsis thaliana.

XX PN WO2003013227-A2.

XX PD 20-FEB-2003.

XX PF 09-AUG-2002; 2002WO-US025805.

XX PR 09-AUG-2001; 2001US-0310847P.

XX PR 19-NOV-2001; 2001US-0336049P.

XX PR 11-DEC-2001; 2001US-0338692P.

XX PR 14-JUN-2002; 2002US-00171468.

XX PA (MEND-) MENDEL BIOTECHNOLOGY INC.

XX PI Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JB;

XX PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;

XX PI Broun PE;

XX DR WPI; 2003-248221/24.

XX DR N-FSDB; ADD30822.

XX PT New plant transcription factor polynucleotides and polypeptides, useful
 PT in producing transgenic plants with commercially valuable properties,
 PT such as an alteration in a plant growth characteristic, e.g. growth rate
 or apomixis.

XX PS Disclosure; SEQ ID NO 852; 454pp; English.

XX CC The invention relates to a number of isolated Arabidopsis thaliana cDNA
 CC sequences and their encoded proteins which are especially transcription
 CC factor related cDNA's and proteins. The isolated or recombinant plant
 CC transcription factor polynucleotides and polypeptides are useful in
 CC producing transgenic plants with commercially valuable properties, i.e.
 CC modified or altered desirable traits as compared to a reference plant,
 CC such as an alteration in a plant growth characteristic, e.g. growth rate,
 CC germination rate of seeds, vigor of plants and seedlings, or leaf and
 CC flower senescence. Sequence information related to the polynucleotides
 CC and polypeptides can also be used in bioinformatic search methods. The
 CC transgenic plant is useful for growing a progeny plant from a parent
 CC plant. This sequence represents one of the proteins of the invention.

XX SQ Sequence 574 AA;

Query Match 33.6%; Score 1047.5; DB 7; Length 574;
 Best Local Similarity 38.7%; Pred. No. 2.8e-80;
 Matches 256; Conservative 83; Mismatches 148; Indels 175; Gaps 20;

QY 2 NNNWLCF-----SLSPYEQNHHRKDVYSSTTTTVDVAGEYCYDPTAASDESAIQTS 54

Db 3 SNNWLGFPSPNNSLPPHEYN----- 24

QY 55 FPSPFQGVVDVDAFTDRDNNSHSRDWDINGCACNNIHNDQDG----PKLENFLGRTTTIYNT 110

Db 25 ----LGLVSDHM--DNPFQTEQNMW-----INPHGGGDEGGEVPKVADFLGVSKPDENQ 73

QY 111 NENVGDSGSGCYGGGDDGGGSLGSLMTKWTURNPV-----DNVDN-----QENGA 158

Db 74 SNELVAYNDSYYPHTN-----SLMPSVQSNVDVVAACDSNTFNSSYHELOESAHN 125

QY 159 AKGLSLSMNSSTSCDNNDSNNVV---AQGKTIDDSVE-----ATPKTTIESFGQ 206

Db 126 LQSLTLMSGTT-----AGNVVDKASPTETGDNASGALAVVETATPRALDTFGQ 177

QY 207 RTSIYRGVTRHRTWGRYEHLWDNSCKREGQTRGRQVYLGSDYKEDKAARSYDLAALKY 266

Db 178 RTSIYRGVTRHRTWGRYEHLWDNSCRREGQSKGRQVYLGSDYKEDKAARSYDLAALKY 237

QY 267 WGTITTTNPFPMSEYEKEVEEMKHMTRQYVVASLRKSGFGRGASIRYGVTRHHQHGSEW 326

Db 238 WGPSTTTNPITNYEKEVEEMKHMTRQEFVAAIRKSGFGRGASIRYGVTRHHQHGSEW 297

QY 327 ARIGRVAGNKDLYLGTGCTGTEBAEAYDIAAKFGLTAFTNFDNRNRYNKAILESPL 386

Db 298 ARIGRVAGNKDLYLGTGCTGTEBAEAYDIAAKFGLTAFTNFDNRNRYNKAILESPL 357

QY 387 I-GSAAKRLKEA-----NR 399

Db 358 IGGGAAKRLKEAQLSSRKREAEIALGSSFYQGGSGTSGTSSRLQIQPYPLSIQ 417

QY 400 PVPMMISNNVSESENGSQNAVQHQVDLSLLHQHQRVNGYNGNLSSEA 459

Db 418 PLEPFLSQNDISHYNNNAHDSSFNHSHYIQ-TQLHLQQTNNYVYVQSSQNSQOLY 476

QY 460 RACFKQEDDQHFLSNTQSLMTNIDHQ--SSVSDSVTVCGNVYGYGQGAAPV-NCD 516

Db 477 NAYLHSPALLHGLVST-SIVDNNNGSGSGSYNTAAFLGN--HGIGSSSTVGSTE 532

QY 517 AYAASEFDYNARNHYFAQQOQTQSGDPFPAAMTNVGNMYHGGEGGEVAPTFTVW 576

Db 533 EPTVTKTDYD-----MPSSDGTGGYSGWTSBVSQGSNPGG-----VFTMW 572

QY 577 ND 578

Db 573 NE 574

RESULT 11

ABR40836
 ID ABR40836 standard; protein; 516 AA.

XX

ABR40836;
 XX DT 16-MAY-2003 (first entry)
 DE Arabidopsis thaliana oil trait related protein sequence SEQ ID NO:428.
 DE Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;
 KW LIP5-like transcription factor caleosin; ATP citrate lyase; SNF1;
 KW CXC-like transcription factor; antisense inhibition; co-suppression;
 KW transgenic plant.
 XX OS Arabidopsis thaliana.
 XX PN WO2003002751-A2.
 XX PD 09-JAN-2003.
 XX PF 27-JUN-2002; 2002WO-US020152.
 XX PR 29-JUN-2001; 2001US-0301913P.
 XX PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX PA (PION-) PIONEER HI-BRED INT INC.
 XX PI Allen SM, Allen WB, Cahoon RE, Epeibaum S, Famodu OO, Harvell LT;
 XX PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
 XX PI Tarczynski MC;
 XX DR WPI; 2003-201509/19.
 XX Novel nucleotide fragment encoding polypeptides having receptor-like
 PT protein kinase activity, caleosin-like activity, useful for altering oil
 PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
 XX Claim 12; Page 469-471; 542pp; English.
 XX The present invention describes an isolated nucleotide fragment (I)
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
 CC activated protein (MAP)-kinase activity, LIP5-like transcription factor
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
 CC activity and CXC-like transcription factor activity. Also described: (1)
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
 CC oil obtained from (V). (I) or its part can be used in antisense
 CC inhibition or co-suppression in a transformed plant. (III) is useful for
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
 CC canola, brassica, sorghum, sunflower or coconut. (III) is also useful for
 CC creating transgenic plants having altered lipid profiles. (I) can also be
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to
 CC ABR40879 represent sequences used in the exemplification of the present
 CC invention
 XX SQ Sequence 516 AA;
 Query Match 33.4%; Score 1040.5; DB 6; Length 516;
 Best Local Similarity 44.6%; Pred. No. 9.3e-80;
 Matches 254; Conservative 70; Mismatches 114; Indels 131; Gaps 22;
 86 NIHNDEODG---PKLENFLGRITTT-IYNTNEN-----VGDGSGSGCYGGGGGSLGL 135
 3 NPHGGGEGGEVPRKVAFLGVSKSGDHTDNLNVPYNDIHQTNASDYFOTN---SL-L 57
 136 SMITWLRNPQVNDVNOENQENAAKGLSLSNSS-----TSCDNNDSN 179
 58 PTVTVCASNAP-NNYELQESAHNLQSLTSLMGSTGAAAEVATVKASPAETSDNSSTT 116
 180 NNVAQAQKTIDDSVEATPKTIESFGORTSIYRGVTRHRTWGRYEAHLWNSCKREGQTR 239
 117 N---TSGGAI---VEATPRRILEIFGORTSIYRGVTRHRTWGRYEAHLWNSCKREGQSR 170

QY 240 KGRVYLGVDYKKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEBEMKMTREQYVASL 299
 DB 171 KGRQ---GGYDKEEKAARAYDLAALKYWGPTTTTTFITNYEKEVEEMKMTREQYVASI 227
 QY 300 RKKSGRSRGASIVRGVTRHGHGRWQARIGRVAGNKDLYLGTGTCGEEAAEAYDIAAIK 359
 DB 228 RKKSGRSRGASIVRGVTRHGHGRWQARIGRVAGNKDLYLGTGTCGEEAAEAYDIAAIK 287
 QY 360 FRGLTAVTNFDMNRVNVKAILSPSLPI-GSAAKRLKEA-----NRVPVPMNMISNNVSE 413
 DB 288 FRGLNAVNTFENRYNDVKAILSENTLPIGGAAKRLKEAQALESRRKREEMIALGNSFHQ 347
 QY 414 SENSASGQNAA-----VOHHGV-----DL\$----- 435
 DB 348 -YGAASGSSSVASSSRLQLQPYPLSIQQPFELHHQPLLTQLQNNNDISQVHDSFSYIQT 406
 QY 436 LLHGHQRYNGYNGNGLSSESARACFKQEDDOHFLSNTQSLMTNIDHQSSVSDSVT 495
 DB 407 QLHLHQOQTNNYLOSSHSTLSQ-----YNAYLQSNPGLL-----HGFVSDNN-- 448
 QY 496 VCGNVVGYGQGPAAFPVNCDAVAASEFDYNARHNYFAQQOQTQOSFGDFFPAAMTNV 555
 DB 449 ---NTSGFLGNNIGIGSSSTVGSSAEEFPAVKVDY-----DMPPS----- 487
 QY 556 GSNMYHGEVGGEVAP-----TFTVWND 578
 DB 488 GGATGYGWNESGESAQSGNPGVFTMNE 516
 RESULT 12
 ABB79640
 ID ABB79640 standard; protein; 585 AA.
 AC ABB79640;
 XX DT 21-OCT-2002 (first entry)
 XX Cotton AINTEGUMENTA-like polypeptide GhANT1.
 DE AINTEGUMENTA; ANT-like polypeptide; GhANT1; transgenic plant; cotton;
 KW plant.
 XX Gossypium hirsutum.
 XX WO200259332-A2.
 XX 01-AUG-2002.
 XX 19-DEC-2001; 2001WO-US049294.
 XX 21-DEC-2000; 2000US-0257896P.
 XX (MONS) MONSANTO TECHNOLOGY LLC.
 XX He SS, Dotson SB;
 XX WPI; 2002-599798/54.
 XX New nucleic acids encoding AINTEGUMENTA-like polypeptides useful in
 PT improving agronomic, horticultural, and quality traits of plants, such as
 PT increased size of plant organs.
 XX Claim 3; Page 158-161; 169pp; English.
 XX The present sequence is the protein sequence of GhANT1, a newly
 CC identified AINTEGUMENTA-like (ANT-like) polypeptide of cotton. The
 CC sequence was predicted from a cDNA clone (see ABN94485) isolated in a
 CC cotton DNA database screening. The invention provides nucleic acids
 CC encoding ANT-like polypeptides comprising, in N-terminal to C-terminal
 CC direction, 2 AP2 DNA binding domains followed by an amino acid
 CC subsequence selected from those given in ABB79629-35. ANT-like
 CC polypeptides were identified in soybean, rice, cotton and corn (see
 CC ABB79636-41). Nucleic acids (see ABN94480-86) encoding the ANT-like

XX 27-JUN-2002; 2002WO-US020152.
XX PF (DUPO) DU PONT DE NEMOURS & CO E I.
XX PR (PION-) PIONEER HI-BRED INT INC.
XX PA
XX PA
XX PI Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;
XX PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;
XX PI Tarczynski MC;
XX DR WPI; 2003-201509/19.
XX DR N-PSDB; ACC00846.
XX PT Novel nucleotide fragment encoding polypeptides having receptor-like
PT protein kinase activity, caleosin-like activity, useful for altering oil
PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.
XX PS
XX PS Claim 12; Page 499-500; 542pp; English.
XX CC The present invention describes an isolated nucleotide fragment (I)
CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a
CC polypeptide (PP) having receptor-like protein kinase activity, mitogen
CC activated protein (MAP)-kinase activity, Lipid-like transcription factor
CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like
CC activity and CKC-like transcription factor activity. Also described: (1)
CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or
CC (II), operably linked to a regulatory sequence; (3) a plant (IV)
CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)
CC oil obtained from (VI). (I) or its part can be used in antisense
CC inhibition or co-suppression in a transformed plant. (III) is useful for
CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,
CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for
CC creating transgenic plants having altered lipid profiles. (I) can also be
CC used as a hybridisation probe. ACC00826 to ACC00868 and ABR40591 to
CC ABR40879 represent sequences used in the exemplification of the present
CC invention
XX SQ Sequence 512 AA;

Query Match 32.0%; Score 999; DB 6; Length 512;
Best Local Similarity 45.9%; Pred. No. 3.2e-76;
Matches 235; Conservative 61; Mismatches 126; Indels 90; Gaps 18;

QY 2 NNNWLGSLSPYEONHHRKDYSSSTTTTVDVAGEYCYDPTAASDESS---AQTSPFSP 58
DB 5 STNWLFSLSLSP-----MDMLRTPPEQFVQ-----YD--AASDTSSHHYLDNLYTNG 49
QY 59 FGVVVDATRONNSHRWDINGCACNNIHDEQD---GPKLENFLGRITTIYNTNEN- 113
DB 50 WG-----NGSLKFEQNLNHSDFVQSSSSVSHAPKLEDFLGDSSAVNRYSDSQ 100
QY 114 --VGDSG-----GSGCYGGGGGGG--GSLGLSMIKTWRNQPDVNDNQENG 156
DB 101 TETQSSSLTHIYDHHHHHHGSSAYFGGDHQLKAITGQAFST-----NSGSEVDDSGASI 156
QY 157 NAAKGLSLSMNSTSCDN-----NNDNNNVVQAQKTIDDSVEAT-----PKYTIES 203
DB 157 GRAQGSSEFGTHIESVNVFAAFSGGTNTGTTLSLAVAQSEKAVAAAESEDRSKKVDDT 216
QY 204 FQORTSIYRGVTRHWTGRYEHLWDNSCKREGQKRGQVYLGKYDKEEKAARAYDLAA 263
DB 217 FQORTSIYRGVTRHWTGRYEHLWDNSCRREGQKRGQVYLGKYDKEEKAARSYDLAA 276
QY 264 LKYWGTTTTNTPMSEYKEVEEMKHMTRQBYVASLRKSSGFSRGASIRGVTRHHQHG 323
DB 277 LKYWGPTATTNPFVSNYKVEVEEMKHVTQKEFIAFLRKSSGFSRGASIRGVTRHHQHG 336
QY 324 RWOARIGRVAGNKDLYLGTGFTQEEAAEAYDIAAIKFRGLTAVTNFDMRYNKAILESP 383
DB 337 RWOARIGRVAGNKDLYLGTGFTATEEAAEAYDIAAIKFRGANAVTNFENRYDVEAIMKS- 395

QY 384 SLPIGSAARLKLK-----EAMRPVPSMMWISNNVSESENSASGWNAAVQH---HQGV-- 432
DB 396 SLPIVGGAARLKLKLSLESEKALPVSS-----SSSSSQOQNPCCGNVSAISINFSSIHQPIAS 451
QY 433 -----DLSLLHQHOERYNGYY-YNGGNLSS 456
DB 452 IPCGIPFDBSTTAYYHNLNFQHFHPTNAGTAAS 483

Search completed: July 15, 2004, 10:01:52
Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 10:00:59 ; Search time 19 Seconds
(without alignments)
1573.234 Million cell updates/sec

Title: US-09-980-364-2

Perfect score: 3119

Sequence: 1 MNNMGLFSLSPYEQNHHRK.....YYHGGGGEVAPFTTWNNDN 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:**

1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCITUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	913.5	29.3	548	4	US-09-227-421-5
2	913.5	29.3	548	4	US-09-479-855-5
3	891	28.6	555	4	US-09-227-421-2
4	891	28.6	555	4	US-09-479-855-2
5	461.5	14.8	432	2	US-08-700-152A-4
6	345	11.1	77	3	US-08-912-272-9
7	345	11.1	77	3	US-09-026-039-9
8	328	10.5	69	3	US-08-912-272-10
9	328	10.5	69	3	US-09-026-039-10
10	188	6.0	67	3	US-08-912-272-11
11	188	6.0	67	4	US-09-026-039-11
12	181	5.8	67	2	US-08-700-152A-1
13	176.5	5.7	68	2	US-08-700-152A-2
14	176.5	5.7	68	3	US-08-912-272-5
15	176.5	5.7	68	4	US-09-026-039-5
16	174	5.6	67	3	US-08-912-272-4
17	174	5.6	67	4	US-09-026-039-4
18	152.5	4.9	375	4	US-09-533-029-12
19	151	4.8	328	4	US-09-964-850-4
20	150	4.8	171	4	US-09-533-029-82
21	150	4.8	248	4	US-09-533-029-22
22	149.5	4.8	328	4	US-09-964-850-6
23	145	4.6	248	4	US-09-202-161B-3
24	144.5	4.6	306	4	US-09-964-850-2
25	141	4.5	264	3	US-09-894-731-4
26	135	4.3	216	4	US-09-138-119C-15
27	135	4.3	216	4	US-09-301-666A-2

28	135	4.3	216	4	US-09-301-217-2	Sequence 2, Appli
29	134.5	4.3	188	4	US-09-198-119C-65	Sequence 65, Appl
30	133	4.3	66	3	US-08-912-272-29	Sequence 29, Appl
31	133	4.3	66	4	US-09-026-039-29	Sequence 26, Appl
32	132.5	4.2	344	4	US-09-533-029-66	Sequence 66, Appl
33	132	4.2	243	4	US-09-533-029-34	Sequence 34, Appl
34	129.5	4.2	68	3	US-08-912-272-28	Sequence 28, Appl
35	129.5	4.2	68	4	US-09-026-039-28	Sequence 28, Appl
36	129	4.1	328	3	US-09-300-672-2	Sequence 2, Appli
37	128.5	4.1	58	4	US-09-202-161B-27	Sequence 27, Appl
38	127.5	4.1	161	4	US-09-202-161B-2	Sequence 2, Appli
39	126.5	4.1	216	4	US-09-198-119C-13	Sequence 13, Appl
40	126.5	4.1	216	4	US-09-301-666A-8	Sequence 8, Appli
41	126.5	4.1	216	4	US-09-301-217-8	Sequence 95, Appl
42	125	4.0	115	4	US-09-198-119C-95	Sequence 14, Appl
43	124.5	4.0	640	4	US-09-336-115C-14	Sequence 2, Appli
44	124	4.0	213	2	US-08-949-603-2	Sequence 2, Appli
45	124	4.0	213	2	US-08-706-270A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-227-421-5
; Sequence 5, Application US/09227421
; Patent No. 6559357
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Mizukami, Yukiko
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility
; TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants
; FILE REFERENCE: 023070-090700PC
; CURRENT APPLICATION NUMBER: US/09/227,421
; CURRENT FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: US 09/227,421
; PRIOR FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; TYPE: PRT
; LENGTH: 548
; ORGANISM: Brassica napus
US-09-227-421-5

Query Match 29.3%; Score 913.5; DB 4; Length 548;
Best Local Similarity 42.0%; Fred. No. 8e-74;
Matches 216; Conservative 57; Mismatches 116; Indels 125; Gaps 13;

QY	4	NMLGFSLSLP-YEQNHHRKDVYSSSTTTTVV-----DVAGEY--CYDPTAASDESS	49
Db	14	NLLGFSLSNNMLKMGGEALYSSSSSVATSSVPPQLVVDNNSNYGVCYGSNLAAREMY	73
QY	50	AIQTSPFPFG---VVDAFTDNNHSHRDWINGCANNIHNDEQDQKLENFLG----	102
Db	74	SQMSVPLRSQSLCLMEALNR--SSHS-----NNHHSQVSPFKMEDFFGTHH	121
QY	103	-----RTTIYNTNENVGDSGCGYGGGGGSLGSLMIKTLRNQ	145
Db	122	NTSHKEMDLSLDFYNTHAPNNTNFOE-----FFSPQT--RNH	162
QY	146	PVDNVDNQEN-----GNAAGLSLSMN--SSTSDNNNDNNNNVVOACK	187
Db	163	HEEETRNVEDPGLTHGGGSFNVGVGFQQSLSLSPGSSCITASHHHQNTQNHQ	222
QY	188	TIDDSVEATP-----KKTIESFGORTSIYRGVTRH	217
Db	223	QISEALVETSAFETTMAAAAKKRGQEVVVGQKQIVHRKSIDTFCORTSIYRGVTRH	282
QY	218	RWTGRYEHLWDSCKREGQTRKGRQVYLGGLGYDKEKAAAYDLAALYKWTGTTTTNPFM	277
Db	283	RWTGRYEHLWDSCKREGQTRKGRQVYLGGLGYDMEKAAAYDLAALYKWTGSPHTNFSV	342

QY 278 SEYEKEVEEMKMTROEYVASLRKSGFSGASIRGVTRHHQHGWRQARIGRVAGNKD 337
 Db 343 ENYQKEIDDMKMTROEYVAHLRKTSGFSGASIRGVTRHHQHGWRQARIGRVAGNKD 402
 QY 338 LYLGTGFTQEEAAEAYDIAIKFRGLTAVTNFDMNRYVKAILESPSPIGSAARLKEA 397
 Db 403 LYLGTGFTQEEAAEAYDIAIKFRGTNAVTFNFDITRYDVRIMASNTLLSGEMARR---- 458
 QY 398 NRPVPSMMIMNNVSESENSASGQNAVAQHOG 431
 Db 459 -----NSNSIVVRNISEDEAALTAUVVNGSGNKEVG 488

RESULT 2
 US-09-479-855-5
 ; Sequence 5, Application US/09479855
 ; Patent No. 639128
 ; GENERAL INFORMATION:
 ; APPLICANT: Fischer, Robert L.
 ; APPLICANT: Mizukami, Yukiko
 ; APPLICANT: The Regents of the University of California
 ; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility
 ; TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants
 ; FILE REFERENCE: 023070-090720US
 ; CURRENT APPLICATION NUMBER: US/09/479,855
 ; CURRENT FILING DATE: 2000-01-07
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 548
 ; TYPE: PRT
 ; ORGANISM: Brassica napus
 ; OTHER INFORMATION: canola AINTEGUMENTA (ANT)
 US-09-479-855-5

Query Match 29.3%; Score 913.5; DB 4; Length 548;
 Best Local Similarity 42.0%; Pred. No. 8e-74;
 Matches 216; Conservative 57; Mismatches 116; Indels 125; Gaps 13;

QY 4 NWLGFSLSPEQYQNHHRKDVYSTTTT-----DVAGEY--CYDPTAASDESS 49
 Db 14 NLLGFSLSNMLKMGGEALYSSSSSSVATSVPPQLVWVGNSSNYGVYCGSNLAAREMY 73
 QY 50 ALOTSFSPSPFG--VVDAFTDRDNNSHSRDWDINGCANNIHNDEQGPKEFLG--- 102
 Db 74 SQMSVPLRSDGSLCLMEALNR--SSHS-----NNHHSQVSPKXNEDFFGPHHH 121
 QY 103 -----RTTITIYNTNENVGDSGCGYGGGGGSLGSLMKTWLRNQ 145
 Db 122 NTSKHEAMDLSLDSLFYNTTHAPNNNTNFQ-----PFSFQT--RNH 162
 QY 146 PVDNVDNQN-----GNAAGLSLNM--SSTSCDNNNSNNVVAQK 187
 Db 163 HEEETRNYENDPGLTHGGSPNVGVYGEFOQSLSLMSPGSQSCITASHHHQNTQNHQ 222
 QY 188 TIDSVSATP-----KKTIESFGQRTSIYRGVTRH 217
 Db 223 QISEALVETSAGFETTTMAAAAKKRGQEVVVGQKQIVHRKSIDTFGQRTSIYRGVTRH 282
 QY 218 RWTGRYEHLWNSCKRGQTRKQROVYLGKYDKEEKAARAYDLAALKYWGTTTTNPFM 277
 Db 283 RWTGRYEHLWNSFKKEGHSRKGQVYLGKYDMEEKAARAYDLAALKYWGPTHTNFVS 342
 QY 278 SEYEKEVEEMKMTROEYVASLRKSGFSGASIRGVTRHHQHGWRQARIGRVAGNKD 337
 Db 343 ENYQKEIDDMKMTROEYVAHLRKTSGFSGASIRGVTRHHQHGWRQARIGRVAGNKD 402
 QY 338 LYLGTGFTQEEAAEAYDIAIKFRGLTAVTNFDMNRYVKAILESPSPIGSAARLKEA 397
 Db 403 LYLGTGFTQEEAAEAYDIAIKFRGTNAVTFNFDITRYDVRIMASNTLLSGEMARR---- 458
 QY 398 NRPVPSMMIMNNVSESENSASGQNAVAQHOG 431

Db 459 -----NSNSIVVRNISEDEAALTAUVVNGSGNKEVG 488

RESULT 3
 US-09-227-421-2
 ; Sequence 2, Application US/09227421
 ; Patent No. 6559357
 ; GENERAL INFORMATION:
 ; APPLICANT: Fischer, Robert L.
 ; APPLICANT: Mizukami, Yukiko
 ; APPLICANT: The Regents of the University of California
 ; TITLE OF INVENTION: Methods for Altering Organ Mass, Controlling Fertility
 ; TITLE OF INVENTION: and Enhancing Asexual Reproduction in Plants
 ; FILE REFERENCE: 023070-090700PC
 ; CURRENT APPLICATION NUMBER: US/09/227,421
 ; CURRENT FILING DATE: 1999-01-08
 ; PRIOR APPLICATION NUMBER: US 09/227,421
 ; PRIOR FILING DATE: 1999-01-08
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 555
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-227-421-2

Query Match 28.6%; Score 891; DB 4; Length 555;
 Best Local Similarity 40.8%; Pred. No. 8.7e-72;
 Matches 227; Conservative 56; Mismatches 122; Indels 152; Gaps 17;

QY 4 NWLGFSLSPEQYQNHHRKDVYSTTTT-----DVAGEY--CYDPTAASDESSA----IQTSF 55
 Db 17 NLLGFSLSNMLKMGGRGREALIYSSSTS-----SAATSSSSVPPQLVWVDN 63
 QY 56 PSPFGV-----VVDAFTDRDNN--HSRD-----WDINGCA 83
 Db 64 TSNFGVCGSNPENGIIYSHMSVPLRSDGSLCLMEALNRSSSHNHQDSSPKVEDFFGTH 123
 QY 84 CNNINDEODGPKLENFLGRITTIYNTNENVD-----GSGSGCYGGGQ--GGG 131
 Db 124 HNNTSHKEAMDLSLDSLFYNTTHEPNTTTFQEPFPPQTRNHEETRYNGNDPSLTHGG 183
 QY 132 SLGLSMIKTWLRNQPVNDVNQENCAAKGJLSLNM--SSTSC-----DNNDS 178
 Db 184 SNVGVY-----GEFOQSLSLMSPGSQSCITGSHHHQONQNHQS 226
 QY 179 NNNVVAQKTTDDSV-----EATPKTIESFGQRTSIYRG 213
 Db 227 QNHQIIEALVETS VGFETTTMAAAKKRGQEDVVVGQKQIVHRKSIDTFGQRTSIYRG 286
 QY 214 VTRHWTGRYEHLWNSCKRGQTRKQROVYLGKYDKEEKAARAYDLAALKYWGTTTTT 273
 Db 287 VTRHWTGRYEHLWNSFKKEGHSRKGQVYLGKYDMEEKAARAYDLAALKYWGPTHT 346
 QY 274 NPFMEYEVEEMKMTROEYVASLRKSGFSGASIRGVTRHHQHGWRQARIGRVA 333
 Db 347 NPSAENYQKEIDDMKMTROEYVAHLRKTSGFSGASIRGVTRHHQHGWRQARIGRVA 406
 QY 334 GNKDIYLTGFTQEEAAEAYDIAIKFRGLTAVTNFDMNRYVKAILESPSPIGSAAR 393
 Db 407 GNKDIYLTGFTQEEAAEAYDIAIKFRGTNAVTFNFDITRYDVRIMASNTLLSGELARR 466
 QY 394 LKEANRPVPSMMIMNN--VSESENSASGQNAVAQHOGVLDLS-----LLH 438
 Db 467 -----NNNSIVVRNTEDQTA--LNAVVEGSGNKEVSTPERLLSPALP 510
 QY 439 QHQRNYGYYNGNLS 455
 Db 511 QVNQMFSGSNM--GGNMS 526

RESULT 4

```

; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,152A
; FILING DATE: 20-AUG-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067200US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-700-152A-4

Query Match 14.8%; Score 461.5; DB 2; Length 432;
Best Local Similarity 34.3%; Pred. No. 3.4e-33;
Matches 114; Conservative 59; Mismatches 112; Indels 47; Gaps 8

QY 198 KXTIESFQRTSIYGRVTRHRWTRGRVEAHLWDMNSCKREGQTKRGQVYLGGVYDKEEKAAR 257
Db 119 KKSRRGPRSSQYRGVTFYRRTRGWESHWD--C-----GKQVILGGFDTAFAAAR 168

QY 258 AYDLAALKYWGCTTTTTFPNSPEYEKEVEEKMKHTROEVVASLRXSSGFSRGASYRGVT 317
Db 169 AYDRAAIKFRGVEADINFDIDDDLKQNTNUTKEEFVHVRSTGPPRGSSKYRGVT 228

QY 318 RHQHRGRQWARGVAGNKDLYLGTGTGTQEEAAEAYDIAAIKFRGLTAVTNFDMNRNVK 377
Db 229 L-HKGRWEARMGQFLGKYYVGLFLDPTVEAARAYDKAAIKCKGKDAVTFNDFPSYDDE 287

QY 378 ALLESFSLPIGSAAKRLKEANRPVPSMMISNNVSESENASGQWAAVQHHQ----- 430
Db 288 LNAES-----SGNPTTPQDNLDSLGNSSANKSHKSODMRMRMNOODSLH 334

QY 431 -----GVDLSLLHQHQRNYNGYNGNGNLSSESARACFKQEDDQHHF--LSNTQSLMTNI 483
Db 335 SNEVLGLCQTGWLNHTNSNHQPPGSSNIGSGGGFSLFPAAEN-HFDFGRASTNQVLNA 393

QY 484 -----DHQSSVSDSDSVTCGVNVGYGQY 507
Db 394 AASSGFSPHHNQIFNSTP HQNWLOTNQFQ 425

RESULT 6
US-08-912-272-9
; Sequence 9, Application US/08912272
; Patent No. 6093874
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP

```

STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/912,272
FILING DATE: 15-AUG-1997
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/879,827
FILING DATE: 20-JUN-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/700,152
FILING DATE: 20-AUG-1996

ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.

REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-067220US

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide

LOCATION: 1..77
OTHER INFORMATION: /note= "ANT-R1 direct repeat"

NAME/KEY: Region
LOCATION: 44..59

OTHER INFORMATION: /note= "putative ANT-R1 amphipathic
OTHER INFORMATION: alpha-helix (SEQ ID NO:37)"

US-08-912-272-9

Query Match 11.1%; Score 345; DB 3; Length 77;
Best Local Similarity 80.5%; Pred. No. 8.3e-24;
Matches 62; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 208 TSIYRGVTRHRTWGRYEAHLWDNSCKREGQTRKGRQVYLGGYDKEEKAARAYDLAALKYW 267
DB 1 TSQYRGVTRHRTWGRYEAHLWDNSFKKEGHSRKGQVYLGGYDMEEKAARAYDLAALKYW 60

QY 268 GTTTTTFPMSEYEKEV 284
DB 61 GPSTHTNFSAEYQKEI 77

RESULT 7
US-09-026-039-9
Sequence 9, Application US/09026039
Patent No. 6329567

GENERAL INFORMATION:
APPLICANT: Jofuku, K. Diane
APPLICANT: Okamuro, Jack K.

TITLE OF INVENTION: Methods for Improving Seeds
NUMBER OF SEQUENCES: 103

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 09/026,039
FILING DATE: 19-FEB-1998
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/912,272
FILING DATE: 15-AUG-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/879,827
FILING DATE: 20-JUN-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/700,152
FILING DATE: 20-AUG-1996

ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.

REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-067230US

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide

LOCATION: 1..77
OTHER INFORMATION: /note= "ANT-R1 direct repeat"

NAME/KEY: Region
LOCATION: 44..59

OTHER INFORMATION: /note= "putative ANT-R1 amphipathic
OTHER INFORMATION: alpha-helix (SEQ ID NO:37)"

US-09-026-039-9

Query Match 11.1%; Score 345; DB 4; Length 77;
Best Local Similarity 80.5%; Pred. No. 8.3e-24;
Matches 62; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 208 TSIYRGVTRHRTWGRYEAHLWDNSCKREGQTRKGRQVYLGGYDKEEKAARAYDLAALKYW 267
DB 1 TSQYRGVTRHRTWGRYEAHLWDNSFKKEGHSRKGQVYLGGYDMEEKAARAYDLAALKYW 60

QY 268 GTTTTTFPMSEYEKEV 284
DB 61 GPSTHTNFSAEYQKEI 77

RESULT 8
US-08-912-272-10
Sequence 10, Application US/08912272
Patent No. 6093874

GENERAL INFORMATION:
APPLICANT: Jofuku, K. Diane
APPLICANT: Okamuro, Jack K.

TITLE OF INVENTION: Methods for Improving Seeds
NUMBER OF SEQUENCES: 103

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834


```

/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/912,272
/ FILING DATE: 15-AUG-1997
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/912,272
/ FILING DATE: 15-AUG-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/879,827
/ FILING DATE: 20-JUN-1997
/ APPLICATION NUMBER: US 08/700,152
/ FILING DATE: 20-AUG-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bastian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 023070-067220US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 69 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY: Peptide
/ LOCATION: 1..69
/ OTHER INFORMATION: /note= "ANT-R2 direct repeat"
/
/ NAME/KEY: Region
/ LOCATION: 37..51
/ OTHER INFORMATION: /note= "putative ANT-R2 amphipathic
/ alpha-helix"
/
/ US-08-912-272-10

```

```

Query Match          10.5%; Score 328; DB 3; Length 69;
Best Local Similarity 91.0%; Pred. No. 2.4e-22;
Matches 61; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 310 ASIYGVTRHHQGRWQARIGRVAGNKDLYLGTGFTQEEAAAYDIAAIKFRGLTAVTNF 369
Db 1 ASIYGVTRHHQGRWQARIGRVAGNKDLYLGTGFTQEEAAAYDIAAIKFRGTNAVTF 60

QY 370 DMNRVNV 376
Db 61 DITRYDV 67

```

```

RESULT 9
US-09-026-039-10
/ Sequence 10, Application US/09026039
/ Patent No. 6329567
/ GENERAL INFORMATION:
/ APPLICANT: Jofuku, K. Diane
/ APPLICANT: Okamuro, Jack K.
/ TITLE OF INVENTION: Methods for Improving Seeds
/ NUMBER OF SEQUENCES: 103
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/026,039
/ FILING DATE: 19-FEB-1998
/ CLASSIFICATION: 800
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/912,272
/ FILING DATE: 15-AUG-1997
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/879,827
/ FILING DATE: 20-JUN-1997
/ APPLICATION NUMBER: US 08/700,152
/ FILING DATE: 20-AUG-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bastian, Kevin L.
/ REGISTRATION NUMBER: 34,774
/ REFERENCE/DOCKET NUMBER: 023070-067230US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 10:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 69 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY: Peptide
/ LOCATION: 1..69
/ OTHER INFORMATION: /note= "ANT-R2 direct repeat"
/
/ NAME/KEY: Region
/ LOCATION: 37..51
/ OTHER INFORMATION: /note= "putative ANT-R2 amphipathic
/ alpha-helix"
/
/ US-09-026-039-10

```

```

Query Match          10.5%; Score 328; DB 4; Length 69;
Best Local Similarity 91.0%; Pred. No. 2.4e-22;
Matches 61; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 310 ASIYGVTRHHQGRWQARIGRVAGNKDLYLGTGFTQEEAAAYDIAAIKFRGLTAVTNF 369
Db 1 ASIYGVTRHHQGRWQARIGRVAGNKDLYLGTGFTQEEAAAYDIAAIKFRGTNAVTF 60

QY 370 DMNRVNV 376
Db 61 DITRYDV 67

```

```

RESULT 10
US-08-912-272-11
/ Sequence 11, Application US/08912272
/ Patent No. 6093874
/ GENERAL INFORMATION:
/ APPLICANT: Jofuku, K. Diane
/ APPLICANT: Okamuro, Jack K.
/ TITLE OF INVENTION: Methods for Improving Seeds
/ NUMBER OF SEQUENCES: 103
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30

```



```

; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067200US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 67 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 1..67
; OTHER INFORMATION: /note= "APTALA 2 (AP2) of Arabidopsis
; OTHER INFORMATION: direct repeat domain AP2-R1 consisting
; OTHER INFORMATION: of amino acids 129-195 of the AP2
; OTHER INFORMATION: protein"
US-08-700-152A-1

Query Match          5.8%; Score 181; DB 2; Length 67;
Best Local Similarity 46.8%; Pred.No.4.2e-09;
Matches 36; Conservative 14; Mismatches 17; Indels 10; Gaps 2;

Qy      208 TSIYRGVTRHWTCRYEAAHLWDNCSCKREGQTRKGVVLGGDYDKEEAARAYDLAALKYW 267
       :|||::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1 SSQRGVTFYFRTRGWESHWD--C-----GKVYLGGFDTAHRAARAAYDRAAKFR 50
       :|||::|||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy      268 GTTTTITNPMPSEYEKEY 284
       |:|||::|||:|||||:|||||:|||||:|||||:|||||:
Db      51 GVPEADINFDIDYDDL 67
       ||:|||::|||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 13
US-08-700-152A-2
; Sequence 2, Application US/08700152A
; Patent No. 5994622
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700.152A
; FILING DATE: 20-AUG-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067200US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Domain
; LOCATION: 1..68
; OTHER INFORMATION: /note= "APETALA 2 (AP2) of Arabidopsis
; OTHER INFORMATION: direct repeat domain AP2-R2 consisting
; OTHER INFORMATION: of amino acids 221-288 of the AP2
; OTHER INFORMATION: protein"
US-08-700-152A-2

Query Match 5.7%; Score 176.5; DB 2; Length 68;
Best Local Similarity 57.6%; Pred. No. 1.1e-08;
Matches 38; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 310 ASIYGVYTHQHGRGHWQARIQGVAGNKLVLGTFGTQEEAAEAYDIAAKFSGLTAVTNF 369
Db 1 SSKYRGVTL-HKCRGWEARMGQFLGKKYVYGLGFDTEVEAARAYDKAAIKCKGKQAVTNF 59
QY 370 DMRYN 375
Db 60 DPSIYD 65

RESULT 14
US-08-912-272-5
; Sequence 5, Application US/08912272
; Patent No. 603874
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,272
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/579,827
; FILING DATE: 20-JUN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/700,152
; FILING DATE: 20-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067220US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..68
; OTHER INFORMATION: /note= "AP2-R2 direct repeat at
; OTHER INFORMATION: positions 221 to 288"

```

```
;
;
; NAME/KEY: Region
; LOCATION: 33..50
; OTHER INFORMATION: /note= "putative AP2-R2 amphipathic
; OTHER INFORMATION: alpha-helix (SEQ ID NO:7)"
US-08-912-272-5

Query Match      5.7%; Score 176.5; DB 3; Length 68;
Best Local Similarity 57.6%; Pred. No. 1.1e-08;
Matches 38; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 310 ASIVRGVTRHHQHGWQRIQGVAGNKDLYLGTGTGTEAAEAYDIAAIKFRGLTAVTNF 369
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 SSKYRGVTL-HKCGRWEARMGQFLGKRYVYLGFLDTEVEARAYDKAAIKNGKDAVTNF 59

QY 370 DMNRYN 375
   :|:|:
Db 60 DPSIYD 65

RESULT 15
US-09-026-039-5
; Sequence 5, Application US/09026039
; Patent No. 6329567
; GENERAL INFORMATION:
; APPLICANT: Jofuku, K. Diane
; APPLICANT: Okamuro, Jack K.
; TITLE OF INVENTION: Methods for Improving Seeds
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026.039
; FILING DATE: 19-FEB-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/912,272
; FILING DATE: 15-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/879,827
; FILING DATE: 20-JUN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/700,152
; FILING DATE: 20-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-067230US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..68
; OTHER INFORMATION: /note= "AP2-R2 direct repeat at
; positions 221 to 288"
```

```
;
;
; NAME/KEY: Region
; LOCATION: 33..50
; OTHER INFORMATION: /note= "putative AP2-R2 amphipathic
; OTHER INFORMATION: alpha-helix (SEQ ID NO:7)"
US-09-026-039-5

Query Match      5.7%; Score 176.5; DB 4; Length 68;
Best Local Similarity 57.6%; Pred. No. 1.1e-08;
Matches 38; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 310 ASIVRGVTRHHQHGWQRIQGVAGNKDLYLGTGTGTEAAEAYDIAAIKFRGLTAVTNF 369
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 SSKYRGVTL-HKCGRWEARMGQFLGKRYVYLGFLDTEVEARAYDKAAIKNGKDAVTNF 59

QY 370 DMNRYN 375
   :|:|:
Db 60 DPSIYD 65

Search completed: July 15, 2004, 10:04:10
Job time : 20 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2004, 10:03:15 ; Search time 49 Seconds
(without alignments)
3693.319 Million cell updates/sec

Title: US-09-980-364-2

Perfect score: 3119

Sequence: 1 MNMNLGSLSPYEQNHKK.....YYHGGGGEVAPFTFTWINDN 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 1285345

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1310	42.0	707	12	US-10-183-687-352
2	1296	41.6	710	12	US-10-183-687-356
3	1131	36.3	692	12	US-10-183-687-530
4	1119.5	35.9	553	12	US-10-183-687-362
5	1111.5	35.6	561	12	US-10-424-599-234179
6	1075.5	34.5	486	12	US-10-183-687-354
7	1047.5	33.6	574	12	US-10-225-066A-356
8	1047.5	33.6	574	12	US-10-225-066A-856
9	1047.5	33.6	574	12	US-10-225-066A-912
10	1047.5	33.6	574	12	US-10-374-780A-334
11	1040.5	33.4	516	15	US-10-183-687-428
12	1011	32.4	585	13	US-10-024-632-11
13	1010.5	32.4	538	12	US-10-183-687-487
14	1003	32.2	470	12	US-10-183-687-485
15	999	32.0	512	12	US-10-183-687-483

16	997.5	32.0	540	12	US-10-183-687-429	Sequence 429, App
17	997	32.0	566	12	US-10-183-687-529	Sequence 529, App
18	994	31.9	557	12	US-10-183-687-491	Sequence 491, App
19	990.5	31.8	510	12	US-10-424-599-243772	Sequence 243772, App
20	986	31.6	510	12	US-10-183-687-348	Sequence 348, App
21	977	31.3	558	12	US-10-225-066A-1014	Sequence 1014, App
22	977	31.3	558	15	US-10-374-780A-2496	Sequence 2496, App
23	974.5	31.2	498	12	US-10-225-066A-382	Sequence 382, App
24	974.5	31.2	498	15	US-10-374-780A-346	Sequence 346, App
25	971.5	31.1	639	12	US-10-437-963-151728	Sequence 151728, App
26	971	31.1	530	12	US-10-183-687-358	Sequence 358, App
27	966	31.0	700	16	US-10-437-963-139609	Sequence 139609, App
28	964.5	30.9	469	12	US-10-183-687-338	Sequence 338, App
29	961	30.8	492	12	US-10-183-687-489	Sequence 489, App
30	960	30.8	557	12	US-10-424-599-213367	Sequence 213367, App
31	957.5	30.7	384	12	US-10-183-687-360	Sequence 360, App
32	956.5	30.7	460	12	US-10-425-114-38886	Sequence 38886, App
33	954.5	30.5	642	13	US-10-024-632-6	Sequence 334, App
34	949.5	30.4	484	12	US-10-183-687-334	Sequence 4793, App
35	948	30.4	410	12	US-10-425-114-40793	Sequence 4793, App
36	947.5	30.4	459	12	US-10-425-114-47823	Sequence 5327, App
37	947	30.4	264	12	US-10-425-114-53727	Sequence 332, App
38	943.5	30.3	489	12	US-10-183-687-332	Sequence 2, Appli
39	942	30.2	663	13	US-10-024-632-9	Sequence 9, Appli
40	942	30.2	669	13	US-10-225-066A-592	Sequence 592, App
41	941	30.2	415	12	US-10-374-780A-400	Sequence 400, App
42	941	30.2	415	16	US-10-437-963-122829	Sequence 122829, App
43	935	30.0	466	16	US-10-425-114-36848	Sequence 36848, App
44	934.5	30.0	354	12	US-10-183-687-421	Sequence 421, App
45	930	29.8	485	12		

ALIGNMENTS

RESULT 1

```
US-10-183-687-352
; Sequence 352 Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu Omclayo O.
; APPLICANT: Harvell, Iselle T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B01458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 352
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Glycine max
US-10-183-687-352
```

Query Match 42.0%; Score 1310; DB 12; Length 707;
Best Local Similarity 45.6%; Pred. No. 3.4e-104;
Matches 352; Conservative 77; Mismatches 139; Indels 168; Gaps 29;

QY

4 NWLGSLSPYEQNHKKVYSSTT-----TTVDVAGEYCYDPTAASDESSAIQ 52

```

; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 356
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Glycine max
US-10-183-687-356

Query Match      41.6%; Score 1296; DB 12; Length 710;
Best Local Similarity 45.5%; Pred. No. 5.5e-103;
Matches 324; Conservative 66; Mismatches 146; Indels 176; Gaps 29;

QY 5 NLLGFSLPFHEHPSQD-HSQTPSRFSPNPDGSISSSTDVAGG-CFDLTS---DSTPHL 59
QY 53 TSFPSPFVVVDATFRDNN-SHSRDWDIN-----GCACN--NIHNBODGPKLENFL 101
Db 60 LNLPS-YG-IYEAHFRNNSINTTDWKENYNSQNLGTSCKNQNNQOQQPKLENFL 117
QY 102 GRTTIYNTENV-GDGGSGGY-----GGDGGG-----GSLGLSMIKTWLR 143
Db 118 GCHS--FGEHQYTCGNASTDYMPAPVAGGSGGGGNNNNNSIGLSMIKTWLR 175
QY 144 NQPDVNDQNGNAAG-----LSLSNV-----SST-----SCDNNDSNN-- 181
Db 176 NQ-PNSENNINNNESGGNRSVQOTLSLWSVTSQSSTSLPLLTASVDNGESSDNKQ 234
QY 182 ---VVAQKTIDDSVTEATPKTIESFGORTSIYRGVTRHRTWGRYEHLWDNSCKREGOT 238
Db 235 PNTSAALDSOTGAIETAPRKSIDTFGORTSIYRGVTRHRTWGRYEHLWDNSCKREGOT 294
QY 239 RKGRQVYLGVDYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKMTQBYVAS 298
Db 295 RKGRQVYLGVDYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKMTQBYVAS 354
QY 299 LERKSSGSRGASIRYGVTRHQRWQARIQVAGNKDLYLGTFTGCEBAEAYDIAAI 358
Db 355 LERKSSGSRGASIRYGVTRHQRWQARIQVAGNKDLYLGTFTGCEBAEAYDIAAI 414
QY 359 KFRGLTAVTNFDMNRYNKAILESPLSPIGSAAKRLKE-----ANRPVPSMM 405
Db 415 KFRGLSANTNDFMSYDVKSILESTTLPIGGAARLKDMQEVLSVDNGHRADQVDSII 474
QY 406 MISNVSESENSASQWNAVOHQGVDSLHLLHQEYRYNGYNGNLSSESARACFKQ 465
Db 475 MSSHLTQGINNYAGGGTAT--HHNWHNAHAFHQPQPCPTTHYPYQORIN-----WCKQE 527
QY 466 EDDQ-----HHFLSNTQSL--MTNIDHQSSVDDSDVTVCNGVV 501
Db 528 QODNSDAPHSLSYSDIHQIQLGNGTHNFFHTNSGLHPMLSMD---SASIDNSSSVV 584
QY 502 --GYGYQGF-AAPV-NCDYAAASEFDYN-----ANR 399
Db 585 YDGYGGGGYVMPGTTTAVVSDGQDQPRSNHGFQDNEIKALGYESVYGSATSDYAH 644
QY 527 ARNHYYFAQQQQTQSPGDFPAAATNNVGSNNYHGGEGGEVAPT 572
Db 645 ARNLYLTQQQ-----SSVDIVKASAYDQGSACNTWVPT 679

```

RESULT 2

```

US-10-183-687-356
; Sequence 356, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: B01458 US NA
; CURRENT APPLICATION NUMBER: US/10/183.687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29

```

RESULT 3

```

US-10-183-687-530
; Sequence 530, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Cahoon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd

```

```

; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 356
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Glycine max
US-10-183-687-356

Query Match      41.6%; Score 1296; DB 12; Length 710;
Best Local Similarity 45.5%; Pred. No. 5.5e-103;
Matches 324; Conservative 66; Mismatches 146; Indels 176; Gaps 29;

QY 4 NMLGFSLSPEY-----QNHR-----KDYISTTTTTVDVAGEYCYDPTAASDESSA 50
Db 5 NLLGFSLSPEHPSQDHSQTSAPRPFNPDGSISS---DVAGD-CFDLTS---DSTP 55
QY 51 ICTSPSPFVVVDATFRDNNSH-SRDWDIN-----GCACN--IHN---DEODGP 95
Db 56 HLNLUPS-YG-IYEAHFRNNSINTTDWKENYNSQNLGTSCKNQNNHROOQOQOOP 113
QY 96 KLENFLGRITTYNTINEN--VGDGSGGY-----GGDGG-----GSLGLSM 137
Db 114 KLENFLGHS--FGEHQYPCGNSASTYMPAPVAGGSGGGGSSNSTSSSIGLSM 171
QY 138 IKTWLRNQPDVNDQNGNAAG-----LSLSNV-----SST-----SCDNNND 177
Db 172 IKTWLRNQPDHSENNNNNNESGNSRSVQOTLSLSNSTGSSQSTSLPLLTASVDNGES 231
QY 178 SNNN-----VVAQKTIDDSVTEATPKTIESFGORTSIYRGVTRHRTWGRYEHLWDNSC 232
Db 232 SDNQKPHHTAALDTTGTGAIETAPRKSIDTFGORTSIYRGVTRHRTWGRYEHLWDNSC 291
QY 233 KREGOTRKGRQVYLGVDYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKMT 292
Db 292 RREGOTRKGRQVYLGVDYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKMT 351
QY 293 QYVASLRKSSGSRGASIRYGVTRHQRWQARIQVAGNKDLYLGTFTGCEBAEAA 352
Db 352 QYVASLRKSSGSRGASIRYGVTRHQRWQARIQVAGNKDLYLGTFTGCEBAEAA 411
QY 353 YDIAAKFRGLTAVTNFDMNRYNKAILESPLSPIGSAAKRLKE-----ANR 399
Db 412 YDVAIAKFRGLSANTNDFMSYDVKSILESTTLPIGGAARLKDMQEVLRVENVHRADQ 471
QY 400 PVPMMNNTNNVSESENSASQWNAVOHQGVDSLHLLHQEYRYNGYNGNLSSES 459
Db 472 EDHSSIMSHLTQGINNYAGGGTAT--HHNWHNHNALAFHQPQPCPTTHYPYQORIN--- 527
QY 460 RACFKQEDDQ-----HHFLSNTQSL--MTNIDHQSSVDDSDVT 495
Db 528 --WCKQEQDSDASHLSYSDIHQIQLGNGTHNFFHTNSGLHPMLSMD---SASIDNSS 582
QY 496 VCGNVV--GYGYQGF-AAPV-NCDYAAASEFDYN----- 526
Db 583 SSNSVYDGYGGGGYVMPGTTTAVVANDGQDQPRSNHGFQDNEIKALGYESVYGT 642
QY 527 -----ARNHYYFAQQQQTQSPGDFPAAATNNVGSNNYHGGEGGEVAPT 572
Db 643 DPHYHAARNLYLTQQQPS-----VDAYKASAYDQGSACNTWVPT 683

```

APPLICANT: Kinney, Tony
APPLICANT: Klein, Ted
APPLICANT: Li, Changjiang
APPLICANT: Oliveira, Igor Cunha
APPLICANT: Sakai, Hajime
APPLICANT: Shen, Bo
APPLICANT: Tarczynski, Mitchell C.
TITLE OF INVENTION: Alteration Of Oil Traits In Plants
FILE REFERENCE: BB1458 US NA
CURRENT APPLICATION NUMBER: US/10/183,687
CURRENT FILING DATE: 2002-06-27
PRIOR FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 532
SOFTWARE: Microsoft Office 97
SEQ ID NO 530
LENGTH: 692
TYPE: PRT
ORGANISM: Oryza sativa
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: gi 2016:013
US-10-183-687-530

Query Match 36.3%; Score 1131; DB 12; Length 692;
Best Local Similarity 40.1%; Pred. No. 1e-88;
Matches 278; Conservative 65; Mismatches 125; Indels 226; Gaps 23;

QY 3 NNWLGFSLSPYEONHHRKDVYSTTTTVDVAGEYCYDPTAADESASAIQTSPPSPFGV 62
Db 5 NNWLAFSLSPQDQLPP-----SQNSTLISAAA-----TTTAGDSTGDVCNIP----- 50
QY 63 VDAFTDRNNSHSDWDINGCACNNIHNDQDQPKLENFLGRITTTIYNTNENVDGQ----- 117
Db 51 -----QDWSRSGLSALVAE-----PKLEDFLGGSFSEQQHHGKGVIPS 94
QY 118 SGSGCY-----GGG-----DG 128
Db 95 SAACVASSSGSVGLYPPSSSSQLQFADSVWATSPVVAHDVGGVGGVMSAAAAAAS 154
QY 129 GGGSLGLSMIKTLRNQFVNDVNDQENGNAAGLSLSMNS----- 168
Db 155 GNGGIGLSMIKWMLRSQAP-----QPAQLSLSMWAGTTTAAQGGAMALLAGAGE 206
QY 169 -----STSC-----DNNDSDNNVVAQKTIIDS-----V 193
Db 207 RGRTPASESLSTSAHGATTATMAGGRKEINEEGSGAGAVAVGSESGSGAVVAGAA 266
QY 194 EATPKTIESFGORTSIYRGVTRHRTGRYEHLWNSCKREGQTRKGRQVYLGVDKEE 253
Db 267 AAARKSVDFGORTSIYRGVTRHRTGRYEHLWNSCRREGQTRKGRQ---GGYDKEE 323
QY 254 KAARAYDLALKYWGTTTTNPFMSYEKEVEEMKHTROEYVASLRKSSGFSRGASII 313
Db 324 KAARAYDLALKYWGTTTTNPFVNNYEKELEEMKHTROEFVASLRKSSGFSRGASII 383
QY 314 RGVTRHHQHGWOARIGRVAGNKDLYLGTGTOEAAEAYDIAAKPRGLTAVTNFDMNR 373
Db 384 RGVTRHHQHGWOARIGRVAGNKDLYLGTGTOEAAEAYDIAAKPRGLNATNFDMSR 443
QY 374 YNVKAILSPSLPGSAAKRLKEANRVP- SMMISNNVS-----ESENSASGW 421
Db 444 YDVKSLDLSAALPVGTAARKLKDAEAAAAYDVGRVASHLGGDGAAYAHYGHHSAAAAA 503
QY 422 QNAAVQ-----HHQGVDSLHLHQHOERYNYNGNLSSESARACPKQEDDHHFLS 474
Db 504 PTIAFAAAAAPPAAAG-----LYHPYAQLRGW-----C-KOEQD-HAVIA 543
QY 475 NTQSLMTNIDHQSSVSDSDSVTCVGNVGYGGYOGFAAPVNCDAVAASEFDYNARNHYIFA 534
Db 544 AAHSIQ-DLHH-----LNLGAAAAA-----HDFFS 567

QY 535 QQQQTQSPGDDFFPAAMTNVGSN-MYYHGEGG 567
Db 568 QAMQOQHGLGSDIDNASLEHSTGNSVYVNGDNGG 601

RESULT 4

US-10-183-687-362
; Sequence 362, Application US/10183687
; Publication No. US20030204870A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Allen, William B.
; APPLICANT: Caloon, Rebecca
; APPLICANT: Epelbaum, Sabine
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Harvell, Leslie T.
; APPLICANT: Jones, Todd
; APPLICANT: Kinney, Tony
; APPLICANT: Klein, Ted
; APPLICANT: Li, Changjiang
; APPLICANT: Oliveira, Igor Cunha
; APPLICANT: Sakai, Hajime
; APPLICANT: Shen, Bo
; APPLICANT: Tarczynski, Mitchell C.
; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
; FILE REFERENCE: BB1458 US NA
; CURRENT APPLICATION NUMBER: US/10/183,687
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: 60/301,913
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 362
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Glycine max
US-10-183-687-362

Query Match 35.9%; Score 1119.5; DB 12; Length 553;
Best Local Similarity 43.5%; Pred. No. 7.2e-88;
Matches 275; Conservative 77; Mismatches 147; Indels 133; Gaps 23;

QY 1 MNNWLGFLSPYEONHHRKDVYSTTTTVDVAGEYCYDPTAADESASAIQTSPPSPFG 60
Db 1 MNNWMLSPFLS-----PTHSLPAHDLQATQYHOPS 31
QY 61 VVDAFTDRNNSHSDWD-INGCACNNIHNDQDQPKLENFLGRITTTIYNTNENVDGSG 119
Db 32 LGLVNNMDNPPQNHNDWNLINTHSSNEI-----PKVADFLG-----VSKENQSDLAA 79
QY 120 SGCYGGGCGGSGLSGLSMIKTLRNQFVNDVND- QENGNA- AKGLSLSMNS- -STSCDN 174
Db 80 LNEIHSNDSLYLFTNNSLVP--MQNPVLDTPSNEYOEANNSNLQSLTTLMSGSGKDJSTCET 137
QY 175 NNDSNNVVAQKTIIDDSVEATPKTIESFGORTSIYRGVTRHRTGRYEHLWNSCKR 234
Db 138 SGENSTNT-----TVEVAPRRLDTFGORTSIYRGVTRHRTGRYEHLWNSCKR 188
QY 235 EGQTRKGRQVYLGVDKEEKAARAYDLALKYWGTTTTNPFMSYEKEVEEMKHTROE 294
Db 189 EGQSRKGRQVYLGVDKEEKAARAYDLALKYWGTTTTNPFISNYKELDEVMKHTROE 248
QY 295 YVASLRKSSGFSRGASIIYRGVTRHHQHGWOARIGRVAGNKDLYLGTGTOEAAEAYD 354
Db 249 FVAAIRRKSSGFSRGASMYRGVTRHHQHGWOARIGRVAGNKDLYLGTGTFSTEEAAEAYD 308
QY 355 IAAIKFRGTLATNFDNMYNVKAILSPSLPI-GSAAKRLKEA-----NRVPVSMMI- 407
Db 309 IAAIKFRGLNATNFDMSRYDVKAILSENTLPIGGAAKRLKEAQALESRRKEEVIAG 368
QY 408 SNNVSESENSASGWQNA--VOHH-----QGVDSLHLH-CHQE---RYNGY 448
Db 369 SSSTFYGTSASSRRLHAYPLAQHHQHPQPOPLLTQLNHDISSSHFSHQDPLHHQGYI 428

QY 449 YNGNLSSESARACPK-QEDDOHH--FLSNTQSL---MTNIDHOSSVS-----DDSVTVVC 497
 Db 429 QTQLQHQQSGGSSYSFQNNAFYNGYLDONHPALLOQMMNGSSSSSVLENNSNNN 488
 QY 498 GNVVGGYGGQGFAPVNCDA-----YAASEFDYNARNHYFAQQOCTOOSPGGD 546
 Db 489 NNVGGFVG-SGFGMASNATAGTGTAEELGLVKVDYDMPAGGYGGRSAADSMQTSNGG- 546
 QY 547 FPAAMTNVGSNMYTHGEGGGEVAPFTVWND 578
 Db 547 -----VFTWMD 553

RESULT 5

US-10-424-599-234179
 ; Sequence 234179, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 234179
 ; LENGTH: 561
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_53488C.1.pap
 US-10-424-599-234179

Query Match 35.6%; Score 1111.5; DB 12; Length 561;
 Best Local Similarity 42.0%; Pred. No. 3.6e-87;
 Matches 273; Conservative 76; Mismatches 140; Indels 161; Gaps 22;

QY 1 MNNWLGFSLGPYEONHHRKDVYSSTTTTVDVAGEYCYDPTAASDESSAIQTSFPPFG 60
 Db 1 MNNWLSPLS-----PHTSLPAHDLQATQYHQS 31
 QY 61 VVDAFTDNNSHSRDWP-INGCANNIHNDQDGPXLENPLGRTTITNTNENVDGDSG 119
 Db 32 LGLVNNNDNPNFQNDWNLINTHSSNEI-----PKVADPLG-----VSKSENQSDLA 79
 QY 120 SGCYGGGGGGLSLGSLMIKTWLNQPDVNDN--QENGNA-AKGLSLSMNS--STSCDN 174
 Db 80 LNEHSDSDYLFNTNSLVP--MNPVLDTPSNYQENANSLQSLTSMGSKDSTCET 137
 QY 175 NDSNNNVVAAOKTIDDSVEATPKTIESFGQRTSIYRGVTRHRWTGRYFAHLWDSCKR 234
 Db 138 SGENSTNT-----TTTTVEVAPRTLDTFGQRTSIYRGVTRHRWTGRYFAHLWDSCKR 192
 QY 235 EGOTKRGQVYLGVDKEEKAARYDLAALKYWGTTTTTTFPMSYEKEVEEKKHMTROE 294
 Db 193 EGQGRKRGQVYLGVDKEEKAARYDLAALKYWGTTTTTTFPMSYEKEVEEKKHMTROE 252
 QY 295 YVASLRKSSGFSRGASIGYRGVTRHHQGRWQARIGEVAGNKDLYLGTFTQBEAAEAYD 354
 Db 253 FVAAIRKSSGFSRGASIGYRGVTRHHQGRWQARIGEVAGNKDLYLGTFTSTEEAAEAYD 312
 QY 355 IAAIKFRGLTAVTFDNNRYNNKAILLESPLPI-GSAAKELKA-----NRPVPSMMNIS 408
 Db 313 IAAIKFRGLNATVTFDNNRYNNKAILLESNTLPIGGGAKEKLAQALLESKRKEEMIALG 372
 QY 409 NVSESENSASGQWNA--VQHH-----QGVDSL-----LHO-----439
 Db 373 SSTFQYGTSSNSRLHAYPLVQCHHHQFEQPOPLTLQNHDISSHFHQDPLHQGYIQTO 432

QY 440 ---HQER-----YNGYYNGNLSSESARACPKQEDDOHHFLSNTQSLMTN 482
 Db 433 LQHQOQSGGSSYSFQNNINNAOYFNGYNL-----ONHPALLOQKINNG 478
 QY 483 IDHQSVDSDSVTVVCNVVYGYGYQGFAPVNC-----CDYAASEFDYNARNHYF 533
 Db 479 SSSSSSVLENNSNNNNTNNVGVFG-SGFGMASNATSGTGTAEELGLVKVDYDMPGGY- 536
 QY 534 AQOQQTQSPGGDFPAA-----MTNVGSNMYTHGEGGGEVAPFTVWND 578
 Db 537 -----CGWSAAAAAESMQTSNG-----VFTWMD 561

RESULT 6

US-10-183-687-354
 ; Sequence 354, Application US/10183687
 ; Publication No. US20030204870A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Cahoon, Rebecca
 ; APPLICANT: Egelbaum, Sabine
 ; APPLICANT: Famodu, Omolayo O.
 ; APPLICANT: Harvell, Leslie T.
 ; APPLICANT: Jones, Todd
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Klein, Ted
 ; APPLICANT: Li, Changjiang
 ; APPLICANT: Oliveira, Igor Cunha
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Shen, Bo
 ; APPLICANT: Tarczyński, Mitchell C.
 ; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
 ; FILE REFERENCE: BBI458 US NA
 ; CURRENT APPLICATION NUMBER: US/10/183,687
 ; CURRENT FILING DATE: 2002-06-27
 ; PRIOR APPLICATION NUMBER: 60/301,913
 ; PRIOR FILING DATE: 2001-06-29
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 354
 ; LENGTH: 486
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; ORGANISM: Glycine max
 US-10-183-687-354

Query Match 34.5%; Score 1075.5; DB 12; Length 486;
 Best Local Similarity 50.2%; Pred. No. 3.8e-84;
 Matches 240; Conservative 45; Mismatches 96; Indels 95; Gaps 12;

QY 167 NSTSCDNNNSNNVAGGKTIDDSVEATPKTIESFGQRTSIYRGVTRHRWTGRYFAH 226
 Db 4 NGESSDNRKQP--NTSALDSTQTGAETAPRKSIDTFQRTSIYRGVTRHRWTGRYFAH 61
 QY 227 LWDNCKKEGQTRKGRQVYLGVDKEEKAARYDLAALKYWGTTTTTTFPMSYEKEVEE 286
 Db 62 LWDNCKKEGQTRKGRQVYLGVDKEEKAARYDLAALKYWGTTTTTTFPMSYEKEVEE 121
 QY 287 MKHMTROEYVASLRKSSGFSRGASIGYRGVTRHHQGRWQARIGEVAGNKDLYLGTFTQ 346
 Db 122 MKHMTROEYVASLRKSSGFSRGASIGYRGVTRHHQGRWQARIGEVAGNKDLYLGTFTQ 181
 QY 347 EBAABAYDIAAKPRGLTAVTFDNNRYNNKAILLESPLSPIGSAAKRLKE-----396
 Db 182 EBAABAYDIAAKPRGLTAVTFDNNRYNNKAILLESPLSPIGSAAKRLKDMQVELSDVN 241
 QY 397 ---ANRPVPSMMNISNVSESENSASGQWNAVOHQGVDSLHLHQHQRVNGYNGGN 453
 Db 242 GHRADQVDHSHIIMSSHLTQGINNVAGGTAT--HHNWHNAFAHQPOCPTTHPYGQR 299
 QY 454 LSSESARACFKQEDDO-----HHFLSNTQSL--MTNIDHOSSV 489
 Db 300 IN-----WCKEQQDNDAPHSLSYSDIHQLOLGNNGTHNFHTNSGLHFLMSWD---SA 351


```
OY 490 SDDSVTCGVNY--GVGYGQGF-AAFV-NCDAAYASEFDYN----- 526
Db 352 SIDNSSNSVYDVGCGGYNWPMGT--TAVASDGDQNPNSNHGFGDNEIKALGYES 411
OY 527 -----ARHYFAQQOQTCQSPGDFPRAATNVGNSNMYYHGGEGVAPT 572
Db 412 VYGSATDSYHAHARNLYLTQOO-----SSVDVTKASAYDOGSACNTWVPT 458

RESULT 7
US-10-225-066A-366
; Sequence 366, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: HEARD, Jacqueline E
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MB10036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 366
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-366

Query Match 33.6%; Score 1047.5; DB 12; Length 574;
Best Local Similarity 38.7%; Pred. No. 1.3e-81;
Matches 256; Conservative 83; Mismatches 148; Indels 175; Gaps 20;

OY 2 NNNWLG-----SLSPYEQNHKRDYVSTTTVVVAVAGEYCDPTAASDESSAIQTS 54
Db 3 SNNWLGFLSPNNSLPPHEYN----- 24
OY 55 FPSFPGVVDAFTDNNKSHSRDWDINGACNHNDEQDG-----PKLENFLGRTTTIYNT 110
Db 25 -----LGLVSDEN--DNFQTEWNN-----INPHGGGDEGGEVPKVADFLGVSKPDENQ 73
OY 111 NENVGDSGSGCYGGGCGGSLGSMIKTWLRNPV-----DNVDN-----QENGNA 158
Db 74 SNHLVAYNDSVYFHTN-----SLMPSVQSNVDVVVAACTNTPNNSYHELQESAHN 125
OY 159 AKGLSLNNSSTSCDNNDSNNVV---AQCETIDDSVE-----ATPKKTIESFGQ 206
Db 126 LQSLTSLMGTT-----AGNVVDVKASPETTGDNASGGALAVVETATPRALDTFGQ 177
OY 207 RTSIYRGVTRHWTRGYEAHLWDNSCKREGQTRKGRQVYLGGYDKKEKAAAPYDLAALKY 266
Db 178 RTSIYRGVTRHWTRGYEAHLWDNSCKREGQTRKGRQVYLGGYDKEDKAARSYDLAALKY 237
```

```
OY 267 WGTTTTTFPMSYEKEVEEMKMTROEYVVASLRKSSGFSRGASIVRGVTRHHQHGRWQ 326
Db 238 WGPSTTTTFITTYEKEVEEMKMTROEYVVAALRRKSSGFSRGASMYRGVTRHHQHGRWQ 297
OY 327 ARIGRVAGNKDLYLGTGTQEEAAEAYDIAAIFKRGITAVTNFDMRYNRYKALLESPL 386
Db 298 ARIGRVAGNKDLYLGTFTSTEEAAEAYDIAAIFKRGILNAVTFEINRYDKALLESSTLP 357
OY 387 I-GSAAKRLKEA-----NR 399
Db 358 IGGAAKRLKEAQALESRRKREAEIALGSSFYQGGSGSTSGSSTSRLLQLOPYPLSIQ 417
OY 400 PVPMSMISNNVSENSASGQNAVAQHGGVDLSLLHQHQRERYNGYNGGNLSSESA 459
Db 418 PLEPFLSLQNDISHYNNNNNAHDSSFNHSHYIQ-TQLHLHQITNNVYLCQSSQNSQO 476
OY 460 RACFKQEDDQHHFLSNTQSLMTNIDHQ--SSVSDSDSVTCGVNVCYGYQGFAPV-NCD 516
Db 477 NAYLHSPALLHGLVST-SIVDNNNNNGSGSYNTAAFLGN---HGIGIGSSSTVGSTE 532
OY 517 AYAASEFDYNARNHYFAQQOQTCQSPGDFPAAATNNVGNVNMYYHGGEGVAPTFTVW 576
Db 533 EFPVTKTDYD-----MPSDGTGGYSGWTSSESVQGSNPGG-----VFTW 572
OY 577 ND 578
Db 573 NE 574

RESULT 8
US-10-225-066A-856
; Sequence 856, Application US/10225066A
; Publication No. US20030226173A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: RATCLIFFE, Oliver
; APPLICANT: RIECHMANN, Jose Luis
; APPLICANT: ADAM, Luc J
; APPLICANT: DUBELL, Arnold T
; APPLICANT: PILGRIM, Marsha L
; APPLICANT: JIANG, Cai-Zhong
; APPLICANT: REUBER, T. Lynne
; APPLICANT: CREELMAN, Robert A
; APPLICANT: PINEDA, Omaira
; APPLICANT: YU, Guo-Liang
; APPLICANT: BROUN, Pierre E
; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MB10036-2 US
; CURRENT APPLICATION NUMBER: US/10/225,066A
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 60/837,444
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 1122
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 856
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-225-066A-856

Query Match 33.6%; Score 1047.5; DB 12; Length 574;
Best Local Similarity 38.7%; Pred. No. 1.3e-81;
Matches 256; Conservative 83; Mismatches 148; Indels 175; Gaps 20;
```

QY 2 NNNWLG-----SLSPYEQNHHRKDVYSTTTTVDVAGEYCYDPTAASDESSAIOTS 54
 Db :|||||
 QY 3 SNNWLGFLPLSPNNSSLPPEYH----- 24
 Db :|||||
 QY 55 FSPSPGVVVDATFRDNNSHRWDINGCACNNIHNDDEODG---PKLENFLGRITTIYNT 110
 Db :|||||
 QY 25 -----LGLVSDHM--DNPFQTQEWNN-----INPHGGGDEGGEVPEKVAADFLGVSKPDENQ 73
 Db :|||||
 QY 111 NENVGDSGGCGYGGGDSGLSLMTKWLNRNPV-----DNVDN-----QENGNA 158
 Db :|||||
 QY 74 SNHLVAYNDSDYFHTN-----SLMPSVQSNVDVVAACDSNTPNNSSVHELQESAHN 125
 Db :|||||
 QY 159 AKGLSLSNMSSFCSDNNDSNNNV-----AQGTIDDSVE-----ATPKTIESFGQ 206
 Db :|||||
 QY 126 LOSLTLSMGT-----AGNVVDKASPETTGDNASGGALAVVETATPRRALDTFGQ 177
 Db :|||||
 QY 207 RTSIYRGVTRHRWTCRYEAHLWDNSCKREGQTRKGRQVYLGKYDKEEKAARAYDLAALKY 266
 Db :|||||
 QY 178 RTSIYRGVTRHRWTCRYEAHLWDNSCRREGQSRKGRQVYLGKYDKEEKAARSYDLAALKY 237
 Db :|||||
 QY 267 MGTITTTNPFMSYKVEVEEMKHTROBYVASLRKSSGFSRGASIIYRGVTRHRHQRWQ 326
 Db :|||||
 QY 238 WGPSTTTNPFITINYEKEVEEMKHTROEFVAAIRKSSGFSRGASMYRGVTRHRHQRWQ 297
 Db :|||||
 QY 327 ARIGRVAGNKDLYLGTFTQTEBEAAEAYDIAAIKFRGLTAVTNFDMNRYNVKAILLESPLP 386
 Db :|||||
 QY 298 ARIGRVAGNKDLYLGTFTSTEEAAEAYDIAAIKFRGLNAVTFINRYDVKAILLESSTLP 357
 Db :|||||
 QY 387 I-GSAAKELKEA----- 399
 Db :|||||
 QY 358 ICGGAATKELKEAQALESRRKREAEWIALGSSFYGGSGSTSGSTRLOLPYPLSIQ 417
 Db :|||||
 QY 400 PVPSSMMWISNNVSESENASQWQAAVQHHCQVDLSLHCHOERYNGYNGNLSSEA 459
 Db :|||||
 QY 418 PLEPFLSLQNNDISHYNNNAHDSSFNHHSYIO-TQLHLHQNTNNYLOQSSONSOLY 476
 Db :|||||
 QY 460 RACFKQEDDQHFLSNTQSLMTNIDHQ--SSVSDSVTVCGNVVGYGQYGFAPV-NCD 516
 Db :|||||
 QY 477 NAYLHSPALLHGLVST-SIVDNNNNGSGSGSYNTAAFLGN---HGIGIGSSSTVGSTE 532
 Db :|||||
 QY 517 AYAASEFDYNARNHYVFAQQOQTOQSGGDFPAAWNTNNGSNMYHGEGGEVAPFTVW 576
 Db :|||||
 QY 533 EFTVKTIDY-----MPSDGTGGYSGWTSVQGSNPGG-----VFTMW 572
 Db :|||||
 QY 577 ND 578
 Db :|||||
 QY 573 NE 574

RESULT 9

US-10-225-066A-912
 ; Sequence 912, Application US/10225066A
 ; Publication No. US20030226173A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendel Biotechnology, Inc.
 ; APPLICANT: RATCLIFFE, Oliver
 ; APPLICANT: RIECHMANN, Jose Luis
 ; APPLICANT: ADAM, Luc J
 ; APPLICANT: DUBELL, Arnold T
 ; APPLICANT: HEARD, Jacqueline E
 ; APPLICANT: PILGRIM, Marsha L
 ; APPLICANT: JIANG, Cai-Zhong
 ; APPLICANT: REUBER, T. Lynne
 ; APPLICANT: CREELMAN, Robert A
 ; APPLICANT: PINEDA, Omaira
 ; APPLICANT: YU, Guo-Liang
 ; APPLICANT: BROWN, Pierre E
 ; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants
 ; FILE REFERENCE: MBI0036-2 US
 ; CURRENT APPLICATION NUMBER: US/10/225,066A
 ; CURRENT FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: 09/837,444

Query Match 33.6%; Score 1047.5; DB 12; Length 574;
 Best Local Similarity 38.7%; Pred. No. 1.3e-81;
 Matches 256; Conservative 83; Mismatches 148; Indels 175; Gaps 20;
 QY 2 NNNWLG-----SLSPYEQNHHRKDVYSTTTTVDVAGEYCYDPTAASDESSAIOTS 54
 Db :|||||
 QY 3 SNNWLGFLPLSPNNSSLPPEYH----- 24
 Db :|||||
 QY 55 FSPSPGVVVDATFRDNNSHRWDINGCACNNIHNDDEODG---PKLENFLGRITTIYNT 110
 Db :|||||
 QY 25 -----LGLVSDHM--DNPFQTQEWNN-----INPHGGGDEGGEVPEKVAADFLGVSKPDENQ 73
 Db :|||||
 QY 111 NENVGDSGGCGYGGGDSGLSLMTKWLNRNPV-----DNVDN-----QENGNA 158
 Db :|||||
 QY 74 SNHLVAYNDSDYFHTN-----SLMPSVQSNVDVVAACDSNTPNNSSVHELQESAHN 125
 Db :|||||
 QY 159 AKGLSLSNMSSFCSDNNDSNNNV-----AQGTIDDSVE-----ATPKTIESFGQ 206
 Db :|||||
 QY 126 LOSLTLSMGT-----AGNVVDKASPETTGDNASGGALAVVETATPRRALDTFGQ 177
 Db :|||||
 QY 207 RTSIYRGVTRHRWTCRYEAHLWDNSCKREGQTRKGRQVYLGKYDKEEKAARAYDLAALKY 266
 Db :|||||
 QY 178 RTSIYRGVTRHRWTCRYEAHLWDNSCRREGQSRKGRQVYLGKYDKEEKAARSYDLAALKY 237
 Db :|||||
 QY 267 MGTITTTNPFMSYKVEVEEMKHTROBYVASLRKSSGFSRGASIIYRGVTRHRHQRWQ 326
 Db :|||||
 QY 238 WGPSTTTNPFITINYEKEVEEMKHTROEFVAAIRKSSGFSRGASMYRGVTRHRHQRWQ 297
 Db :|||||
 QY 327 ARIGRVAGNKDLYLGTFTQTEBEAAEAYDIAAIKFRGLTAVTNFDMNRYNVKAILLESPLP 386
 Db :|||||
 QY 298 ARIGRVAGNKDLYLGTFTSTEEAAEAYDIAAIKFRGLNAVTFINRYDVKAILLESSTLP 357
 Db :|||||
 QY 387 I-GSAAKELKEA----- 399
 Db :|||||
 QY 358 ICGGAATKELKEAQALESRRKREAEWIALGSSFYGGSGSTSGSTRLOLPYPLSIQ 417
 Db :|||||
 QY 400 PVPSSMMWISNNVSESENASQWQAAVQHHCQVDLSLHCHOERYNGYNGNLSSEA 459
 Db :|||||
 QY 418 PLEPFLSLQNNDISHYNNNAHDSSFNHHSYIO-TQLHLHQNTNNYLOQSSONSOLY 476
 Db :|||||
 QY 460 RACFKQEDDQHFLSNTQSLMTNIDHQ--SSVSDSVTVCGNVVGYGQYGFAPV-NCD 516
 Db :|||||
 QY 477 NAYLHSPALLHGLVST-SIVDNNNNGSGSGSYNTAAFLGN---HGIGIGSSSTVGSTE 532
 Db :|||||
 QY 517 AYAASEFDYNARNHYVFAQQOQTOQSGGDFPAAWNTNNGSNMYHGEGGEVAPFTVW 576
 Db :|||||
 QY 533 EFTVKTIDY-----MPSDGTGGYSGWTSVQGSNPGG-----VFTMW 572
 Db :|||||
 QY 577 ND 578
 Db :|||||
 QY 573 NE 574

RESULT 10

US-10-374-780A-334
 ; Sequence 334, Application US/10374780A

Publication No. US20040019927A1
GENERAL INFORMATION:
APPLICANT: Sherman, Bradley K
APPLICANT: Riechmann, Jose Luis
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Heard, Jacqueline E
APPLICANT: Haake, Volker
APPLICANT: Creelman, Robert A
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc J
APPLICANT: Reuber, T. Lynne
APPLICANT: Keddle, James
APPLICANT: Brown, Pierre E
APPLICANT: Pilgrim, Marsha L
APPLICANT: Dubell III, Arnold T
APPLICANT: Pineda, Omaira
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
FILE REFERENCE: WBI-0047 CIP
CURRENT APPLICATION NUMBER: US/10/374,780A
PRIOR FILING DATE: 2003-02-25
CURRENT FILING DATE: 2003-02-25
PRIOR FILING DATE: 09/837,944
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/934,455
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 10/225,066
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,067
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,068
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 2906
SOFTWARE: Patent in version 3.2
SEQ ID NO 334
LENGTH: 574
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: GI793 (conserved domain in AA coordinates:179-255, 281-349)
US-10-374-780A-334

Query Match 33.6%; Score 1047.5; DB 15; Length 574;
Best Local Similarity 38.7%; Pred. No. 1.3e-81;
Matches 256; Conservative 83; Mismatches 148; Indels 175; Gaps 20;
Qy 2 NNNWLGPF-----SLSPYEQNHKRDVYSTTTTVDVAGEYCYDPTAASDESSAIQTS 54
Db 3 SNNWLGFLSPNNLSLPPHEYN-----24
Qy 55 FPSFPGVVDAFTDRNNSHSDWDINGCACNNIHNDQDG-----PKLENFLGRITTTIYNT 110
Db 25 ----LGLVSDHM--DNPPQTEWNM-----INPHGGGDEGEVYPKVADFLGVSKPDENQ 73
Qy 111 NENVGDSGGSCYGGGGGSLGSMKTKWLRQPV-----DNVDN-----QENGNA 158
Db 74 SNHLVAYNDSYFYHTN-----SLMPVQSNVDVVAACTSNTPNNSSVHELQESAHN 125
Qy 159 AKGLSLMNSSTSCDNNNDNNNVV---AOGKTIIDDSVE-----ATPKTIESFGQ 206
Db 126 LQSLTSLMGTT-----AGNVVDKASPTETGDNASGALAVVETATPRRALDIFGQ 177
Qy 207 RTSIYRGVTRHRTWTRYEAHLWDNSCKREGQTRKGRQVYLGVDYKEKAAPAYDLAALKY 266
Db 178 RTSIYRGVTRHRTWTRYEAHLWDNSCKREGQTRKGRQVYLGVDYKEDKAARSYDLAALKY 237

Qy 267 WGTITTTNFMSEYEKEVEEMKHMTRQEVYASLRRKSSGSRGASLYRGVTRHHQGRWQ 326
Db 238 WGPSTTTNFPITNYEKEVEEMKHMTRQEFVAAIRKSSGSRGASMYRGVTRHHQGRWQ 297
Qy 327 ARIGRVAGNKDLYLGTGTQOEAAAEAYDIAAIAKFERGLTAVTNFDMNRYNKAILESPSLP 386
Db 298 ARIGRVAGNKDLYLGTSTEBEAAAEAYDIAAIAKFERGLNAVTFNFEINVDVKAILLESSTLP 357
Qy 387 I-GSAAKRLKEA-----NR 399
Db 358 ICGGAARLKEAQALESRRKEAEMIALGSSFYGGGSGTSGSTSSRLQLQPYPLSIQQ 417
Qy 400 PVPSSMMISNNVSESENSAGWQAAVQHOGVDSLHLHQHQERYNGYYNGNLSBSA 459
Db 418 PLEPFLSLQNDIDISHYNNNNHADSSFNHHSYIQ-TQLHLHQQTNNYLQQOSSQNSQQLY 476
Qy 460 RACFKQEDDQHFLSNTQSLMTNIDHQ--SSVSDSDSVTCNVGVYGYQGFAPV-NCD 516
Db 477 NAYLHSPALLHLGLVST-SIVDNNNNNGSSGYNTAAFLGN---HGIIGSSSTVSGSTE 532
Qy 517 AYAASEFDYNARNHYFFAQOQOQTQOSPGGDFPAAATNNVGSNNVYHGBGGGEVAPTFTVW 576
Db 533 EFTVTKTIDY-----MPSSDGTGGYSGWTSVQGSNPGG-----VFTMW 572
Qy 577 ND 578
Db 573 NE 574
RESULT 11
US-10-183-687-428
Sequence 428, Application US/10183687
Publication No. US20030204870A1
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Allen, William B.
APPLICANT: Cahoon, Rebecca
APPLICANT: Epeibaum, Sabine
APPLICANT: Famodu, Omolayo O.
APPLICANT: Harvell, Leslie T.
APPLICANT: Jones, Todd
APPLICANT: Kinney, Tony
APPLICANT: Klein, Ted
APPLICANT: Li, Changjiang
APPLICANT: Oliveira, Igor Cunha
APPLICANT: Sakai, Hajime
APPLICANT: Shen, Bo
APPLICANT: Tarczynski, Mitchell C.
TITLE OF INVENTION: Alteration Of Oil Traits In Plants
FILE REFERENCE: BB1458 US NA
CURRENT APPLICATION NUMBER: US/10/183,687
CURRENT FILING DATE: 2002-06-27
PRIOR APPLICATION NUMBER: 60/301,913
PRIOR FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 532
SOFTWARE: Microsoft Office 97
SEQ ID NO 428
LENGTH: 516
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: gi 4836931
US-10-183-687-428
Query Match 33.4%; Score 1040.5; DB 12; Length 516;
Best Local Similarity 44.6%; Pred. No. 4.5e-81;
Matches 254; Conservative 70; Mismatches 114; Indels 131; Gaps 22;
Qy 86 NIHNDQDG---PKLENFLGRITTT-IYNTNEN-----VGDGSGSCYGGGCGGSLGL 135
Db 3 NPHGGGGEVGPVKADFLGVSKSGDHTDHLVPYNDIHOITNADSYFQTN-----SL-L 57

Qy	136	SMIKTWLRNQVDNVNDONGENNAKGLSLSNSS-----TSCDNNDSN	179
Db	58	PTVVCASNAP--NNVELQESAHN:QSLTSLWGSTGAAAAEVAIVVKASPAETSA DNSSTT	116
Qy	180	NNVVAOGKTI DSVBATPKKTIESFGORTSIYRGVTRHRWTGRYEAHLWDNSCKREGQTR	239
Db	117	N--TSGGAI--VEATPRTLETFGORTSIYRGVTRHRWTGRYEAHLWDNSCRREGQSR	170
Qy	240	KGRQVYLGGYDKEEAAAYDLAALKYWGTTTTTNFPMSEYKEVEEMKHTTROEVAASL	299
Db	171	KGRQ---GGYDKEEAAAYDLAALKYWGSTTTNFPIYKEVEEMKHTTROEVAASI	227
Qy	300	RRKSGFSRGASIIYRGVTRHHQGRWQARIGRVAGNKDLYLGTFTQTEBAABAYDIAIK	359
Db	228	RRKSGFSRGASMYRGVTRHHQGRWQARIGRVAGNKDLYLGTFTSEBAABAYDIAIK	287
Qy	360	FRGLTAVTNFDMNRVNVZAILESPLPI-GSAARLKEA---NRPVPSMMVIMNNVSE	413
Db	288	FRGLNAVTNFENRYDVVKAILESNTLPIGGGAARLKEAQLBESSRKREMTALCSNFHQ	347
Qy	414	SENSASGWNAR-----VQHCGV-----DLS-----	435
Db	348	-YGAASGSSVASSSRLOQYPPLSIQOPEHLHHQPOLTLQNNNDISQYHDSFSYIQT	406
Qy	436	LJHQRQERYNGYINGGNLSSESARACFKQEDDQHFLPNTQSLMTNIDHQSVDSDSVT	495
Db	407	QLHLHQOQTNNYLQSSSHTSQL-----YNAVLOQNPGLL-----HGFVSDNN--	448
Qy	496	VCGNVVGYYGQYGFAPVNCDAYAASEFDYNARHHYFAQQOQTQOSPGCDPFAAMTNV	555
Db	449	----NTSGFLGNNGTIGSSSTVGSSEEPFAVKVDY-----DMPPS-----	487
Qy	556	GSNMYHGGGGEVAP-----TTFVWND	578
Db	488	GCATGYGWNNGESAQGSNPGCVFTMWE	516

RESULT 12
 US-10-024-632-11
 ; Sequence 11, Application US/10024632
 ; Publication NO. US20020170093A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Monsanto Technology LLC
 ; APPLICANT: He, Steve S.
 ; APPLICANT: Dotson, Stanton B.
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ASSOCIATED WITH PLANT CELL PROLIFERATION AND
 ; TITLE OF INVENTION: GROWTH AND USES THEREOF
 ; FILE REFERENCE: 38-21(51,837)B
 ; CURRENT APPLICATION NUMBER: US/10/024,632
 ; CURRENT FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: US 60/257,896
 ; PRIOR FILING DATE: 2000-12-21
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 11
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: Gossypium hirsutum
 US-10-024-632-11

```

109 QY -----NT-----NNVGDGSGGCVGGD----- 127
94 Db GTHENQNTVPQPTRINVNVAYSS-----SGDAEAAENFTNPSSFQTVRYNNYNENPQTLM 149
128 QY GGGGSL-----GLSMIKTWFLENQVDN-----VDNOENEN-AA 159
150 Db AGGSHLQOCCDPNPNHNQBSGVHHVFPFESATSVSGFKSWLRQTFFPGGKASAGNETNNFNF 209
160 QY KGLSLSMKSSTSCDNNDSNNN--VVAQGTKIDD-----SVEATPKKTIESPQGT 208
210 Db QALSUTMSP-----SRNGFFAIAPLEVDNWRKRPVGNKLTRESVPRKSIDTFQGT 261
209 QY SIYRGVTRHRWTGRYEAHLMWNSCKRGQTRKGRQVYLGDYKDEEKAAPAYDLAALKYWG 268
262 Db SQYRGVTRHRWTGRYEAHLMWNSCKRGQTRKGRQVYLGDYKDEEKAAPAYDLAALKYWG 321
269 QY TTTTTFNPKSYEKEVEEMKHWTROEVASLRKSSGFRGASIRYGVTRHHQHGRWOAR 328
322 Db PTHINFPSTIYKELSEMKNMTQEEVAHLRRKSSGFRGASVYRGVTRHHQHGRWOAR 381
329 QY IGRVAGNKDLYLGTGTQEEAAEAYDIAAKFRGLTAVTNFDMNRVNVKAILSPSLPIG 388
382 Db IGRVAGNKDLYLGTFTQEEAAEAYDIAAKFRGTSAVTNFDISRYDVKRICSSSTLIGG 441
389 QY SAAKELKEANEPVPSMMMI SNVYSESENSAGHQAAVOHQHGVDLSLLLHOQERYNGY 448
442 Db ELAKK-----SPKDTASIAEDYNSCASSASPOPLAIP----- 475
449 QY YNGGNLSSESARACFKQEDQHHLFSNTOQLMTNIDHQSSVSDDSVTCGNVVGYYGQG 508
476 Db ---SGEASDELADMWYNTANSDEQQ-----QHQSNTNNDSLANSSSRSSNPQSPKGSIG 528
509 QY FAAPVNCDAVAASEFFDYNAENHXYEACQOQCTQOSPGCDFPAAMTNVNSNMYYHGBGGGE 568
529 Db LAS-----DKFGIGG-DYS--HHGYFSLKGSYEDGNGSETDNSNENRNLGNLGLVH----- 575
569 QY VAPTFTVWMD 578
576 Db KIPFALWNE 585

```

RESULT 13
US-10-183-687-487
/ Sequence 487, Application US/10183687
/ Publication No. US20030204870A1
/ GENERAL INFORMATION:
/ APPLICANT: Allen, Steve
/ APPLICANT: Allen, William B.
/ APPLICANT: Cahoon, Rebecca
/ APPLICANT: Eperlbaum, Sabine
/ APPLICANT: Famodu, Omolayo O.
/ APPLICANT: Harvell, Leslie T.
/ APPLICANT: Jones, Todd
/ APPLICANT: Kinney, Tony
/ APPLICANT: Klein, Ted
/ APPLICANT: Li, Changjiang
/ APPLICANT: Oliveira, Igor Cunha
/ APPLICANT: Sakai, Hajime
/ APPLICANT: Shen, Bo
/ APPLICANT: Tarczynski, Mitchell C.
/ TITLE OF INVENTION: Alteration Of Oil Traits In Plants
/ FILE REFERENCE: BBI458 US NA
/ CURRENT APPLICATION NUMBER: US/10/183.687
/ CURRENT FILING DATE: 2002-06-27
/ PRIOR APPLICATION NUMBER: 60/301,913
/ PRIOR FILING DATE: 2001-06-29
/ NUMBER OF SEQ ID NOS: 532
/ SOFTWARE: Microsoft Office 97
/ SEQ ID NO 487
/ LENGTH: 538
/ TYPE: PRT
/ ORGANISM: *Cossypium hirsutum*
US-10-183-687-487

Query Match 32.4%; Score 1010.5; DB 12; Length 538;
 Best Local Similarity 43.9%; Pred. No. 1.9e-78;
 Matches 249; Conservative 51; Mismatches 138; Indels 129; Gaps 18;

QY 3 NNWLGLSLP--YEQHHRKQVYSSTTTVDVAGYCYDPTAASDESSAIOISPPSPG 60
 DB 7 NNSLAFSLSHHHQNH--LSFFSSSNPTLFOA-----FNTSAA-----AAVPTTTPPA-- 54
 QY 61 VVDAFTRDNNSHRDWDINGCACNNIHNDQD-----GPKLENFLGRTTITYNTNENV 114
 DB 55 -----NLSAEKADAGGEPATTASATTDLMSFKGPKLEDFLGSSIT----- 97
 QY 115 GDGSGCYGGGGGSLGLSMIKTWLRNPVDVNDQENGNAAGLS-----LSMNSSTS 171
 DB 98 -----PGVGAPQLQLHHHNFSTQTPVTVDNDIYSELKTIAAFLRGFSSEQ 147
 QY 172 CDNNDSNNVAAQKTTIDSVBATEPKKTTESFGQRTSIYRGVTRHRWTRGYEAHLWDNS 231
 DB 148 IDIHKQOQOAPIA-----EPAPKESVDTFGQRTSIYRGVTRHRWTRGYEAHLWDNS 198
 QY 232 CKREGQTRKRGQVYLGYYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKMT 291
 DB 199 CRREGQSRKRGQVYLGYYDKEEKAARAYDLAALKYWGTTTTTTFPISNYEKELEEMKMT 258
 QY 292 ROEVASLRKSGFSGASIRGVTRHHQHWQARIGRVAGNKDLYLGTFTQOEAAE 351
 DB 259 ROEVASLRKSGFSGASIRGVTRHHQHWQARIGRVAGNKDLYLGTFTQOEAAE 318
 QY 352 AYDIAAKPGLTAVTNFDMNRYNKAILESPLPSGSAKRLKEANRPVSMNMV 411
 DB 319 AYDIAAKPGLTAVTNFDMNRYNKAILESPLPSGSAKRLKEANRPVSMNMV 371
 QY 412 SESENSASG-----WQNAVQHHQGVLDLSL-----LH 438
 DB 372 TDEERHSGSGMTFTSQPASSTVGFPAINKQDPDSYWSNVLGYNDTGASLSAKNPIF 431
 QY 439 QHBERYNGYVY-----NGEN--LSSESARACFK-----QEDDOHFLSNQSLMT-- 481
 DB 432 HSPNGSGFGSPSGFCTGNGNSAVNGSINGLFNRGYGVQEQQSSAATNSHPLATPI 491
 QY 482 ---NIDHQSSVDDSVTVCGNVGYGG 505
 DB 492 VLANSNDYESSSS-----GYGG 508

RESULT 14

US-10-183-687-485
 ; Sequence 485, Application US/10183687
 ; Publication No. US20030204870A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Allen, William B.
 ; APPLICANT: Cahoon, Rebecca
 ; APPLICANT: Epelbaum, Sabine
 ; APPLICANT: Famodu, Omolayo O.
 ; APPLICANT: Harvell, Leslie T.
 ; APPLICANT: Jones, Todd
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Klein, Ted
 ; APPLICANT: Li, Changjiang
 ; APPLICANT: Oliveira, Igor Cunha
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Shen, Bo
 ; APPLICANT: Tarczyński, Mitchell C.
 ; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
 ; FILE REFERENCE: B01458 US NA
 ; CURRENT APPLICATION NUMBER: US/10/183,687
 ; CURRENT FILING DATE: 2002-06-27
 ; PRIOR APPLICATION NUMBER: 60/301,913
 ; PRIOR FILING DATE: 2001-06-29
 ; NUMBER OF SEQ ID NOS: 532
 ; SOFTWARE: Microsoft Office 97

; SEQ ID NO 485
 ; LENGTH: 470
 ; TYPE: PRT
 ; ORGANISM: Catalpa sp.
 US-10-183-687-485

Query Match 32.2%; Score 1003; DB 12; Length 470;
 Best Local Similarity 45.7%; Pred. No. 6.8e-78;
 Matches 235; Conservative 49; Mismatches 94; Indels 136; Gaps 16;

QY 2 NNNMLGFLSLPYEQNHHRKQVYSSTTTVDVAGYCYDPTAASDESSAIOISPPSPG 61
 DB 7 HQWLAFSLSL-----NHS-----LFHPSFNPNSSSC-----GV 35
 QY 62 VVDAFTRDNNSHRDWDINGCACNNIHND--EQDGPKELENFLGRTTITYNTNENVGDSG 119
 DB 36 V-----NQENG-----AGGAATDHLHLSVLASGPKLEDFLGASC-----GSS 73
 QY 120 SGCYGGGGGGGSLGLSMIKTWLRNPVDVNDQENGNAAGLSLSMNSSTS CDNNDSN 179
 DB 74 GGHITQYDGAQGVNISDVEIY-----GDSELKTIAAFLRGLSHERSDDSQ 120
 QY 180 NNVAQKTTIDSVBATEPKKTTESFGQRTSIYRGVTRHRWTRGYEAHLWDNSCKREGQTR 239
 DB 121 KQTAVAAP-----EPPAKKAAETFGQRTSIYRGVTRHRWTRGYEAHLWDNSCRREGQSR 175
 QY 240 KGRQVYLGYYDKEEKAARAYDLAALKYWGTTTTTTFPMSEYEKEVEEMKMTROEYVASL 299
 DB 176 KGRQVYLGYYDKEEKAARAYDLAALKYWGTTTTTTFPVSDYEREMEEMKMTROEYVASL 235
 QY 300 RKSSGFSRGASIRGVTRHHQHWQARIGRVAGNKDLYLGTFTQOEAAEAYDIAAK 359
 DB 236 RKSSGFSRGASIRGVTRHHQHWQARIGRVAGNKDLYLGTFTQOEAAEAYDIAAK 295
 QY 360 FGLTAVTNFDMNRYNKAILESPLPSGSAKRLKEANRPV-----SMMNISNV 411
 DB 296 FGLTAVTNFDMNRYNKAILESPLPSGSAKRLKEANRPV-----SMMNISNV 354
 QY 412 S-ESENSAS-----GWQNAVQHHQGVLDLSL-----HQQVDLSLH----- 438
 DB 355 SFPSPNSAAAINFALPIKQDHSLSLWALGYQNSITNTNTPNTSVSLFHDNSFMBFQ 414
 QY 439 -----QHBERYNGYVYCGNVLSSES 458
 DB 415 ANNEGFFNGCGYFQHQE-----NGTNTSSTS 441

RESULT 15

US-10-183-687-483
 ; Sequence 483, Application US/10183687
 ; Publication No. US20030204870A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Allen, William B.
 ; APPLICANT: Cahoon, Rebecca
 ; APPLICANT: Epelbaum, Sabine
 ; APPLICANT: Famodu, Omolayo O.
 ; APPLICANT: Harvell, Leslie T.
 ; APPLICANT: Jones, Todd
 ; APPLICANT: Kinney, Tony
 ; APPLICANT: Klein, Ted
 ; APPLICANT: Li, Changjiang
 ; APPLICANT: Oliveira, Igor Cunha
 ; APPLICANT: Sakai, Hajime
 ; APPLICANT: Shen, Bo
 ; APPLICANT: Tarczyński, Mitchell C.
 ; TITLE OF INVENTION: Alteration Of Oil Traits In Plants
 ; FILE REFERENCE: B01458 US NA
 ; CURRENT APPLICATION NUMBER: US/10/183,687
 ; CURRENT FILING DATE: 2002-06-27
 ; PRIOR APPLICATION NUMBER: 60/301,913
 ; PRIOR FILING DATE: 2001-06-29
 ; NUMBER OF SEQ ID NOS: 532

